CUNY2013

Columbia, SC

March 21-23, 2013 — 26th Annual CUNY Conference on Human Sentence Processing

UNIVERSITY OF SOUTH CAROLINA®

Columbia Metropolitan Convention Center

1101 Lincoln Street · Columbia, SC

Columbia Metropolitan Convention Center Guide for Network Users

For **FREE** Hot Spot *wireless* Internet access

(2Mbps bandwidth shared among ALL free wireless users)

- Hotspot wireless SSID = **CMCC_Hotspot**
- Guest access login is required
- For Guest access, please use your <u>complete</u> email address (ex.: johndoe@mycompany.com)

Open your browser with an 802.11a/b/g/n compliant device and you will see the log in screen shown below. If your machine does not automatically redirect to the screen below, manually insert **https://bluesocket.columbiacvb.com/login.pl** into your address bar.

Fill in the appropriate box (Don't forget to check the 'I accept terms…' box) Then click on 'Log In'.

"word-clouded brain" cover art by tagxedo.com

Contents

Program Committee

Fernanda Ferreira, Amit Almor, Dirk-Bart den Ouden

Reviewers

Amit Almor, Jennifer Arnold, Sudha Arunachalam, Markus Bader, Moises Betancort, Allison Blodgett, Ina Bornkessel-Schlesewsky, Dianne Bradley, Holly Branigan, Mara Breen, Richard Breheny, Sarah Brown-Schmidt, Marc Brysbaert, Helen Cairns, David Caplan, Katy Carlson, Craig Chambers, Franklin Chang, Kiel Christianson, Harald Clahsen, Charles Clifton, Derya Cokal, Martin Corley, Seana Coulson, Wind Cowles, Matthew Crocker, Dirk-Bart den Ouden, Michael Walsh Dickey, Wouter Duyck, Thomas Farmer, Yasmeen Faroqi-Shah, Evelina Fedorenko, Fernanda Ferreira, Victor Ferreira, Janet Fodor, Stephani Foraker, Julie Franck, Jerid Francom, Lyn Frazier, Steven Frisson, Susanne Gahl, Alan Garnham, Susan Garnsey, Carlos Gelormini, Silvia Gennari, Edward Gibson, Peter Gordon, Dan Grodner, Martin Hackl, Robert Hartsuiker, Daphna Heller, Barbara Hemforth, Roberto Heredia, Masako Hirotani, Yi Ting Huang, E. Matthew Husband, Erika Hussey, José Manuel Igoa, Kiwako Ito, Scott Jackson, Florian Jaeger, Elsi Kaiser, Yuki Kamide, Andrew Kehler, Laura Kertz, Albert Kim, Robert Kluender, Lars Konieczny, Ellen Lau, Kerry Ledoux, Roger Levy, Richard Lewis, Simon Liversedge, Maryellen MacDonald, Alec Marantz, Andrea E. Martin, Brian McElree, Katherine Messenger, Don Mitchell, Nicola Molinaro, Rebecca Nappa, Shukhan Ng, Mante Nieuwland, Akira Omaki, Pat O'Seaghdha, Nikole Patson, Neal Pearlmutter, Thomas Pechmann, Colin Phillips, Martin Pickering, Maria Pinango, Maria Polinsky, Lucia Pozzan, Liina Pylkkänen, Hugh Rabagliati, Keith Rayner, Hannah Rohde, Douglas Roland, Jeffrey Runner, Amy Schafer, Christoph Scheepers, Matthias Schlesewsky, Petra Schumacher, Carson Schutze, Florian Schwarz, Irina Sekerina, Shari Speer, Adrian Staub, Linnaea Stockall, Britta Stolterfoht, Patrick Sturt, Benjamin Swets, Whitney Tabor, Michael Tanenhaus, Matt Traxler, Annie Tremblay, John Trueswell, Mieko Ueno, Virginia Valian, Julie Van Dyke, Justine VanDyke-Lyon, Shravan Vasishth, Constanze Vorwerg, Matt Wagers, Michael Wagner, Tessa Warren, Tom Wasow, Duane Watson, Thomas Weskott, Ming Xiang, Hiroko Yamashita, Eiling Yee, Masaya Yoshida

Thank You to the Conference Sponsors

College of Arts and Sciences, University of South Carolina

Department of Communication Sciences and Disorders, University of South Carolina

Department of English, Morehead State University

Department of Linguistics, Stanford University

Department of Linguistics, University of Chicago

Department of Linguistics, University of Maryland

Department of Linguistics, University of Massachusetts, Amherst

Department of Psychology, Harvard University

Department of Psychology, New York University

Department of Psychology, UC San Diego

Department of Psychology, University of South Carolina

Departments of Psychology and Linguistics, University of Michigan

The Institute for Mind and Brain, University of South Carolina

The Language and Brain Lab, Carleton University

The Linguistics Program, CUNY Graduate Center

The Linguistics Program, University of South Carolina

National Science Foundation

Neuroscience of Language Lab, NYU and NYU Abu Dhabi

SR Research

And Thanks to Our Generous Donors

Ted Gibson (Junior CUNY angel), Florian Jaeger (Junior CUNY angel), Hiroko Yamashita (Junior CUNY angel), Craig Chambers (patron), Brian W. Dillon (patron), E. Matthew Husband (patron), Randi Martin (patron), Jeffrey T. Runner (patron), Michael K. Tanenhaus (patron), Jennifer E. Arnold, Inbal Arnon, Mara Breen, Katy Carlson, Uriel Cohen Priva, Michael Walsh Dickey, Evelina Fedorenko, Susanne Gahl, Jana Häussler, Elsi Kaiser, Laura Kertz, Maryellen MacDonald, Gail Mauner, Robert Slevc, Patrick Sturt, Titus von der Malsburg, Tessa Warren, Duane Watson

Jerrold J. Katz Young Scholar Award

Named in memory of our friend and distinguished colleague, the Jerrold J. Katz Young Scholar Award recognizes the paper or poster presented at the Annual CUNY Conference on Human Sentence Processing that best exhibits the qualities of intellectual rigor, creativity, and independence of thought exemplified in Professor Katz's life and work. Any first author of a presentation, who is pre-doctoral or up to three years post-PhD and not yet tenured, is eligible for consideration. The amount of the award is $500.

Previous Recipients

Sol Lago and Wing Yee Chow (University of Maryland, College Park), jointly, for their paper entitled "Word frequency affects pronouns and antecedents identically: Distributional evidence", presented at the 24th Annual CUNY Conference on Human Sentence Processing, Palo Alto, CA, March 2011.

Adriana Hanulíková (Max Planck Institute for Psycholinguistics) for her paper entitled "When grammatical errors do not matter: An ERP study on the effect of foreign-accent on syntactic processing", presented at the 23rd Annual CUNY Conference on Human Sentence Processing, New York, NY, March 2010.

Adrian Staub (University of Massachusetts, Amherst) for his paper entitled "The timing of garden path effects on eye movements: Structural and lexical factors", presented at the 22nd Annual CUNY Conference on Human Sentence Processing, Davis, CA, March 2009.

Gunnar Jacob (University of Dundee) for his paper entitled "An inter-lingual garden path? L1 interference in L2 syntactic processing", presented at the 21st Annual CUNY Conference on Human Sentence Processing, Chapel Hill, NC, March 2008.

T. Florian Jaeger (University of Rochester) and Neal Snider (Stanford University), jointly, for their paper entitled "Implicit learning and syntactic persistence: Surprisal and cumulativity", presented at the 20th Annual CUNY Conference on Human Sentence Processing, La Jolla, CA, March 2007.

Scott Jackson (University of Arizona), for his paper entitled "Prosody and logical scope in English", presented at the 19th Annual CUNY Conference on Human Sentence Processing, New York, NY, March 2006.

Sachiko Aoshima (American University), for her paper entitled "The source of the bias for longer filler-gap dependencies in Japanese", presented at the 18th Annual CUNY Conference on Human Sentence Processing, Tucson, AZ, March–April 2005.

Andrew Nevins (Massachusetts Institute of Technology), for his paper entitled "Syntactic and semantic predictors of tense: An ERP investigation of Hindi", presented at the 17th Annual CUNY Conference on Human Sentence Processing, College Park, MD, March 2004.

Britta Stolterfoht (Max Planck Institute of Cognitive Neuroscience), for her poster entitled "The difference between the processing of implicit prosody and focus structure during reading: Evidence from brain-related potentials," presented at the 16th Annual CUNY Conference on Human Sentence Processing, Cambridge, MA, March 2003.

John Hale (Johns Hopkins University), for his paper entitled "The information conveyed by words in sentences," presented at the 15th Annual CUNY Conference on Human Sentence Processing, New York, NY, March 2002.

Award Fund - To make a contribution to the Jerrold J. Katz Fund, please send a check made out to "CUNY Graduate Center (with the notation "Jerrold J. Katz Fund" in the memo line) to: Diane C. Bradley (Katz Award Fund), Ph.D. Program in Linguistics, CUNY Graduate Center, 365 Fifth Avenue, New York, NY, 10016-4309.

Special Session

Theories of Sentence Processing and the Neuroscience of Language

The cognitively based theories that are influential in psycholinguistics today are largely based on non-neural data and theoretical constructs. Theories of sentence processing are still largely based on formal linguistic models of grammar as well as cognitive models of memory, attention, and learning. Ultimately, however, given that language processing must take place in a physical structure, our goal must be to develop theories that are biologically plausible and compatible with other theories in the cognitive neurosciences. This brings us to the fundamental question we want our speakers to address in the special session: Does the basic architecture of language developed in the 1950s and 1960s based primarily on linguistic evidence, or in the 1980s and 1990s based on statistical constraint based models, survive an era of brain imaging, brain stimulation, and sophisticated cognitive neuropsychology? If not, how should we carve up the language system based on what we have learned from the entire range of relevant evidence, including linguistic, behavioral, and biological? Our view is that now is the time to revisit the entire architecture of the language system and to ask whether the modules, architectures, and processing systems that have been assumed up to now need to be profoundly revised in light of what is known about language and the brain. The special session will bring together six prominent researchers with diverse backgrounds to consider this basic issue. All have extensive experience working on the neuroscience of language using a range of methods and techniques, all are major figures in the field of language who represent different theoretical perspectives, and all have made major contributions to the literature on language processing. Yet they all have unique and complementary expertise in the neuroscience of language, which makes them uniquely qualified to begin this challenging but important conversation.

Invited Speakers

Evelina Fedorenko is a research scientist in the Department of Brain & Cognitive Sciences at MIT. She seeks to understand the nature of the representations and computations that underlie language understanding and production, focusing especially on brain localization for different language functions. She is also interested in the relationship between the language system and other cognitive/neural systems.

Julius Fridriksson is a Professor in the Department of Communication Science and Disorders at the University of South Carolina. His research focuses on understanding speech comprehension and production in normal and disorder populations, particularly those who have suffered stroke. His research relies on technologies such as Magnetic Resonance Imaging and transcranial Direct Current Stimulation (tDCS).

Peter Hagoort is director of the Max Planck Institute for Psycholinguistics (since November 2006), the founding director of the Donders Centre for Cognitive Neuroimaging (1999), and a professor in cognitive neuroscience at the Radboud University Nijmegen. His research interests relate to the domain of the human language faculty and how it is instantiated in the brain. In his research he applies neuroimaging techniques such as ERP, MEG, PET and fMRI to investigate the language system and its impairments as in healthy adults as well as in conditions such as aphasia, dyslexia and autism.

Gina R Kuperberg is an Associate Professor in the Dept. of Psychology at Tufts University and an Associate Psychiatrist in the Dept. of Psychiatry at Massachusetts General Hospital. Her lab focuses on the cognitive neuroscience of thought and language in healthy individuals and in those with psychiatric disorders. She uses ERPs, MEG, and fMRI to study both the temporal and spatial dimensions of cognition in the brain.
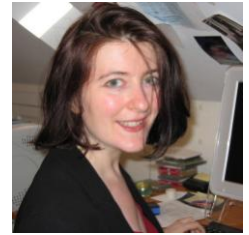
Liina Pylkkänen is an Associate Professor of Linguistics and Psychology at New York University. Her research aims to characterize the representational and processing properties of the combinatory system that supports linguistic creativity. Her research primarily makes use of MEG, which offers the best combination of temporal and spatial resolution among currently available cognitive neuroscience methods.

Mark Seidenberg is Professor of Psychology at the University of Wisconsin, Investigator in the Communication and Cognitive Processes Unit of the Waisman Center, University of Wisconsin, and Senior Scientist at Haskins Labs, New Haven CT. Mark studies language and reading, with the goal of understanding how these skills are acquired and used, and the brain circuits that support them. The work involves a combination of behavioral studies, neuroimaging, and computational (connectionist) modeling.

Thursday Program

8:15 – 9:00	Registration, Coffee, and Light Breakfast	Pre-Ballroom
9:00 – 11:00	Session 1	Ballrooms A & B

9:00	**Welcoming remarks.** *Anne Bezuidenhout* (University of South Carolina)	
9:15	**Should psycholinguistics ignore the language of the brain?** *Peter Hagoort* (Max Planck Institute for Psycholinguistics)	35
10:00	**Sentence type effects in Granger causality analysis of MEG/EEG signal.** *David Caplan, David Gow, Reid Vancelette, Alexander Conrad Nied* (MGH)	36
10:30	**The P600 indexes rational error correction within a noisy-channel model of human communication.** *Edward Gibson[1], Laura Stearns[2], Leon Bergen[1], Marianna Eddy[1], Evelina Fedorenko[1]* ([1]MIT, [2]Wellesley College)	37

11:00 – 11:30	Coffee break
11:30 – 1:00	Session 2

11:30	**To a man with a hammer everything looks like a nail: How explicit lexical predictions influence sentence processing – an ERP study.** *Jakub Szewczyk[1] and Herbert Schriefers[2]* ([1]Jagiellonian University, [2]Radboud University)	38
12:00	**Exploring representations of event duration in language.** *Gitte Joergensen and Silvia Gennari* (University of York)	39
12:30	**Bi-directional structural priming between mathematics and language.** *Christoph Scheepers[1] and Patrick Sturt[2]* ([1]University of Glasgow, [2]University of Edinburgh)	40

1:00 - 2:15	Lunch and Eyetracking Workshop
2:15 – 4:00	Session 3

2:15	**Building meanings in theory vs. in the brain.** *Liina Pylkkänen* (New York University)	41
3:00	**Teasing apart coercion and surprisal: Evidence from ERPs and eye-movements.** *Francesca Delogu, Heiner Drenhaus, Matthew Crocker* (Saarland University)	42
3:30	**The experiments that we finished: Structural separation reduces the cost of coercion.** *Matthew W. Lowder and Peter C. Gordon* (University of North Carolina at Chapel Hill)	43

4:00 – 4:30	Coffee break
4:30 – 6:30	Session 4

Thursday Program

4:30	**Going the distance: Pronoun resolution relies on direct-access retrieval from memory.** *Stephani Foraker and Bryan Wight* (SUNY College at Buffalo)	44
5:00	**Individual differences in anaphoric processing: Insights from mouse-tracking.** *Elsi Kaiser and Alexis Harper* (University of Southern California)	45
5:30	**Contributions of declarative memory to on-line reference resolution: Findings from amnesia.** *Sarah Brown-Schmidt[1], Jake Kurczek[2], Melissa Duff[2]* ([1]University of Illinois, [2]University of Iowa)	46
6:00	**When timing is (almost) everything: Referential dynamics in parent-child interactions.** *John Trueswell[1], Yi Lin[1], Erica Cartmill[2], Benjamin Armstrong[1], Susan Goldin-Meadow[2], Lila Gleitman[1]* ([1]Univ. of Pennsylvania, [2]Univ. of Chicago)	47
6:30 - 8:30	Poster Session 1	

Friday Program

8:15 – 9:00	Late registration, coffee, and light breakfast	Pre-Ballroom
9:00 – 10:45	Session 5	Ballrooms A & B

9:00	**Predicting meaning: What the brain tells us about the architecture of language comprehension.** *Gina Kuperberg* (MGH/Tufts University)	115
9:45	**Language processing in schizophrenia: Top-down & bottom-up effects.** *Hugh Rabagliati[1], Nate Delaney-Busch[2], Jesse Snedeker[1], Gina Kuperberg[2]* ([1]Harvard University, [2]Tufts University)	116
10:15	**Predicting the *foreseeable future*: MEG evidence for preactivation of predicted words.** *Tal Linzen, Joseph Fruchter, Masha Westerlund, Alec Marantz* (New York University)	117

10:45 – 11:15	Coffee break
11:15 – 12:45	Session 6

11:15	**Interference in covert dependencies.** *Ming Xiang[1], Yanling Cui[2], Suiping Wang[2]* ([1]University of Chicago, [2]South China Normal University)	118
11:45	**Trainability and selective transferability of conflict resolution skills to parsing and non-parsing domains.** *Erika Hussey, Susan Teubner-Rhodes, Alan Mishler, Isaiah Harbison, Jared Novick* (University of Maryland)	119
12:15	**The (un)automaticity of syntactic processing in language production: Structural priming is disrupted by verbal memory load.** *Iva Ivanova, Liane Wardlow Lane, Tamar Gollan, Victor Ferreira* (UCSD)	120

12:45 - 2:45	Lunch and Poster Session 2
2:45 - 4:30	Session 7

2:45	**Domain-specific vs. domain-general mechanisms in language learning and processing.** *Evelina Fedorenko* (MIT)	121
3:30	**Individual differences in verbal working memory predict reanalysis vs. integration difficulty in syntax-semantics conflict scenarios.** *Leif Oines, Albert Kim, Akira Miyake* (University of Colorado Boulder)	122
4:00	**Incremental parsing, gapping, and connectives.** *Masaya Yoshida[1], Katy Carlson[2], Michael Walsh Dickey[3]* ([1]Northwestern University, [2]Morehead State University, [3]University of Pittsburgh)	123

4:30 – 5:00	Coffee break
5:00 – 7:15	Session 8

Friday Program

5:00	Conference Organizers, Various	
5:15	**Word order affects the time-course of sentence formulation in Tzeltal.** *Elisabeth Norcliffe, Agnieszka E. Konopka, Penelope Brown, Stephen C. Levinson* (Max Planck Institute for Psycholinguistics)	124
5:45	**The role of interactivity on cognitive alignment and decision making during dialogue.** *Moreno I. Coco[1], Rick Dale[2], Frank Keller[1]* ([1]University of Edinburgh, [2]University of California (Merced))	125
6:15	**Motor plans and linguistic perspective in action sentences: A causal role in comprehension.** *Madeleine Beveridge[1], Daniel Casasanto[2], Roberto Bottini[2], Martin Pickering[1]* ([1]University of Edinburgh, [2]New School for Social Research, New York)	126
6:45	**Perspective taking in online language processing.** *Xiaobei Zheng and Richard Breheny* (University of College London)	127

Saturday Program

8:15 – 9:00	Late registration, coffee, and light breakfast	Pre-Ballroom
9:00 – 10:45	Session 9	Ballrooms A & B

9:00	**Patient studies of language in the modern era.** *Julius Fridriksson* (University of South Carolina)	195
9:45	**Sentential context modulates early phases of visual word recognition: Evidence from a training manipulation.** *Vicky Lai[1,2], Albert Kim[3], James McQueen[1,2,4]* ([1]Max Planck Institute for Psycholinguistics; [2]Donders Institute for Brain, Cognition and Behaviour, [3]University of Colorado Boulder, [3]Radboud University)	196
10:15	**Partially activated words participate in combinatory semantic interpretation during sentence processing.** *Sarah Johnstone, John Trueswell, Delphine Dahan* (University of Pennsylvania)	197

10:45 – 11:15	Coffee break
11:15 – 12:45	Session 10

11:15	**Syntactic priming in comprehension: Priming 'early' closure.** *Matt Traxler, Megan Boudewyn, Tamara Swaab* (UC Davis)	198
11:45	**Local coherence and digging-in effects in German.** *Dario Paape,Titus von der Malsburg, Shravan Vasishth,* (University of Potsdam)	199
12:15	**Task effects on prosodic prominence.** *Andrés Buxó-Lugo, Joe Toscano, Duane Watson* (University of Illinois at Urbana-Champaign)	200

12:45 - 2:45	Lunch and Poster Session 3
2:45 - 4:30	Session 11

2:45	**I remember connectionism.** *Mark Seidenberg* (University of Wisconsin-Madison)	201
3:30	**Grammatical constraints on phonological encoding in speech production.** *Jordana Heller and Matthew Goldrick* (Northwestern University)	202
4:00	**More than words: The effect of multi-word frequency and constituency on phonetic duration.** *Inbal Arnon[1] and Uriel Cohen Priva[2]* ([1]University of Haifa, [2]Brown University)	203

4:30 – 5:00	Coffee break
5:00 – 7:00	Session 12

Saturday Program

5:00	**Syntactic adaptation: Converging on the statistics of the linguistic environment.** *Alex Fine[1], Thomas Farmer[2], T. Florian Jaeger[1]* ([1]University of Rochester, [2]University of Iowa)	204
5:30	**Direct experience versus abstract knowledge in linguistic processing.** *Emily Morgan and Roger Levy* (UCSD)	205
6:00	**A rational inference approach to aphasic language comprehension.** *Edward Gibson[1], Chaleece Sandberg[2], Evelina Fedorenko[1], Swathi Kiran[2]* ([1]MIT, [2]Boston University)	206
6:30	**Comprehension and acquisition of contrastive prosody: Rational inference helps adults and children cope with noisy input.** *Chigusa Kurumada[1], Meredith Brown[2], Michael Tanenhaus[2]* ([1]Stanford University, [2]University of Rochester)	207

1.1	**Processing of novel compounds in adults and children: One word or two?** *Yuki Hirose[1], Takefumi Ohki[1], Reiko Mazuka[2]* ([1]The University of Tokyo, [2]RIKEN Brain Science Institute)	51
1.2	**Thee, uhh, role of discourse status in three-year-olds' understanding of disfluent utterances.** *Sarah Owens and Susan Graham* (University of Calgary)	52
1.3	**Processing effects on grammar acquisition: Evidence from an artificial language study.** *Lucia Pozzan, Lila Gleitman, John Trueswell* (University of Pennsylvania)	53
1.4	**Can anaphoric dependencies be primed across languages? Evidence from Italian-English bilinguals.** *Emily Fedele, Elsi Kaiser, Maria Luisa Zubizarreta* (University of Southern California)	54
1.5	**Children's and adults' processing of noun phrase conjunctions: An eye-tracking study.** *Justine VanDyke-Lyon[1], Lap-Ching Keung[2], Fernanda Ferreira[1]* ([1]University of South Carolina, [2]University of North Carolina, Chapel Hill)	55
1.6	**The online processing of the Japanese anaphoric expressions zibun-zisin and kare.** *Atsushi Yuhaku[1] and Satoru Nakai[2]* ([1]Ritsumeikan University, [2]Doshisha University)	56
1.7	**Resolving temporary referential ambiguity using presupposed content.** *Jacopo Romoli[1], Manizeh Khan[2], Yasutada Sudo[3], Jesse Snedeker[2]* ([1]Macquarie University, [2]Harvard University, [3]Institut Jean-Nicod, CNRS/ENS)	57
1.8	**Coherence expectations underlie parallelism effects for conjoined clauses.** *Laura Kertz and Corey Cusimano* (Brown University)	58
1.9	**Effects of event-structure and topic/focus-marking on pronoun reference in Korean.** *Kitaek Kim, Theres Grüter, Amy Schafer* (University of Hawai'i)	59
1.10	**Clause structure matters: The role of left dislocations and clefts in pronoun resolution.** *Barbara Hemforth[1], Israel de la Fuente Velasco[2], Saveria Colonna[3], Sarah Schimke[4]* ([1]LLF, CNRS, Paris Diderot, [2]LLF, CNRS, Paris Diderot, Labex EFL, [3]SFL, CNRS, Paris 8, [4]University of Osnabrück)	60
1.11	**Online sensitivity to structural constraints on bound variable anaphora.** *Ian Cunnings[1], Clare Patterson[2], Claudia Felser[2]* ([1]University of Edinburgh, [2]University of Potsdam)	61
1.12	**Syntactic prominence in the processing of reference: Does subordination matter?** *Wei Cheng, Jenn Olejarczyk, Amit Almor* (University of South Carolina)	62
1.13	**Figuring out Kafka: Structural biases induce early sense commitment for metonyms.** *Joel Fishbein and Jesse Harris* (Pomona College)	63
1.14	**Walking the walk and talking the talk, and perceptually simulating both while reading.** *Mallory Stites and Kiel Christianson* (University of Illinois, Urbana-Champaign)	64

1.29	**The hypothetical property of "if"-statements: A visual-world paradigm eye-tracking study.** *Likan Zhan, Stephen Crain, Peng Zhou* (Macquarie University)	79
1.30	**Implicatures in uncooperative contexts: Evidence from a visual world paradigm.** *Anna Pryslopska* (Eberhard Karls Universität Tübingen, SFB 833)	80
1.31	**Focus inhibits free associates.** *Mary Byram Washburn, Elsi Kaiser, Maria Luisa Zubizarreta* (University of Southern California)	81
1.32	**Incremental computation of scalar implicatures: An ERP study.** *Les Sikos, Sam Tomlinson, Hilary Traut, Daniel Grodner* (Swarthmore College)	82
1.33	**Stress position congruency hinders word production: Evidence from the picture-word interference paradigm.** *Claudio Mulatti[1], Simone Sulpizio[2], Remo Job[2]* ([1]University of Padua, [2]University of Trento)	83
1.34	**Lexical differentiation in language production and comprehension.** *Si On Yoon and Sarah Brown-Schmidt* (University of Illinois, Urbana-Champaign)	84
1.35	**Does message similarity facilitate sentence formulation?** *Agnieszka Konopka[1], Stefanie Kuchinsky[2], Antje Meyer[3]* ([1]Max Planck Institute for Psycholinguistics; Donders Institute for Brain, Cognition, and Behavior, [2]Medical University of South Carolina, [3]Max Planck Institute for Psycholinguistics; Radboud Universiteit Nijmegen)	85
1.36	**Incremental planning of complex noun phrases in sentence production.** *Maureen Gillespie[1], Victor S. Ferreira[2], T. Florian Jaeger[3]* ([1]University of New Hampshire, [2]University of California San Diego, [3]University of Rochester)	86
1.37	**Modeling word duration in language production.** *Andrés Buxó-Lugo, Dominique Simmons, Duane Watson* (University of Illinois at Urbana-Champaign)	87
1.38	**How do speakers think for speaking in a VOS language?** *Takuya Kubo[1], Manami Sato[1], Hajime Ono[2], Hiromu Sakai[1]* ([1]Hiroshima University, [2]Kinki University)	88
1.39	**Comparing measures of word confusability and their effect on speech production.** *Esteban Buz and T. Florian Jaeger* (University of Rochester)	89
1.40	**Structure selection during sentence production: A role for executive control?** *Maartje van de Velde[1], Agnieszka E. Konopka[1], Antje S. Meyer[2]* ([1]MPI for Psycholinguistics, Nijmegen, [2]MPI for Psycholinguistics and Radboud University, Nijmegen)	90
1.41	**Effects of animacy on processing relative clauses in older and younger adults.** *Gayle DeDe* (University of Arizona)	91
1.42	**Effects of syntactic complexity in an incremental sentence/sentence dual task.** *Joshua Levy[1] and William Evans[2]* ([1]University of Massachusetts Amherst, [2]Boston University)	92

1.43   **Verbal WM capacities in sentence comprehension: Evidence from aphasia.** *Yingying Tan[1], Randi Martin[1], Julie Van Dyke[2]* ([1]Rice University, [2]Haskins Laboratories)   93

1.44   **Modeling individual differences in processing deficits in aphasia.** *Umesh Patil, Sandra Hanne, Shravan Vasishth, Frank Burchert* (University of Potsdam)   94

1.45   **Parasitic gaps inside subject islands in (non-)native sentence processing: Evidence from eye movements during reading.** *Oliver Boxell and Claudia Felser* (University of Potsdam)   95

1.46   **Aspectual interpretation and increment size: A cross-linguistic eyetracking study.** *Oliver Bott and Anja Gattnar* (SFB 833, Tübingen University)   96

1.47   **Dissociating reanalysis and semantic reinterpretation during garden-path recovery.** *Gunnar Jacob and Claudia Felser* (Potsdam Research Institute for Multilingualism, University of Postdam)   97

1.48   **Filler complexity in wh-extractions from islands and non-islands.** *Constantin Freitag and Sophie Repp* (Humboldt-Universität zu Berlin)   98

1.49   **Agreement violations in Arabic: Qualitative ERP differences between singular and plural subjects.** *R. Muralikrishnan[1] and Ali Idrissi[2]* ([1]New York University Abu Dhabi, [2]UAE University Al Ain)   99

1.50   **Acceptability of grammatical and ungrammatical doubly nested relative clause structures in Spanish: some evidence in favor of usage-based approaches.** *Florencia Reali* (Universidad de los Andes)   100

1.51   **Retrieval respects crossover.** *Dave Kush, Colin Phillips, Jeff Lidz* (University of Maryland, College Park)   101

1.52   **The interpretation of elided reflexives in children and adults.** *Sharese King[1] and Jeffrey Runner[2]* ([1]Stanford University, [2]University of Rochester)   102

1.53   **Effects of 'long-before-short' on processing of canonical and scrambled order in Japanese.** *Katsuo Tamaoka[1], Chi Yui Leung[1], Sachiko Kiyama[2]* ([1]Nagoya University, [2]National Center for Geriatrics and Gerontology)   103

1.54   **Resumption rescues islands after all: An experimental investigation of Italian and English.** *Andrea Beltrama and Ming Xiang* (University of Chicago)   104

1.55   **A connectionist model of Mandarin relative clause processing asymmetries.** *Yaling Hsiao and Maryellen MacDonald* (University of Wisconsin-Madison)   105

1.56   **Effects of verb meaning on lexical integration in agrammatic aphasia.** *Jennifer Mack, Woohyuk Ji, Cynthia Thompson* (Northwestern University)   106

1.57 **Similarity-based interference is required for the LIFG effect of object extraction: Evidence from MEG.** *Kimberly Leiken and Liina Pylkkänen* (New York University) 107

1.58 **MEG evidence for neural mechanisms in the reading of Chinese compounds.** *Chun-Hsien Hsu and Chia-Ying Lee* (Academia Sinica) 108

1.59 **What does the left prefrontal cortex do for sentence production? Evidence from tDCS.** *Nazbanou Nozari[1], Jennifer Arnold[2], Sharon Thompson-Schill[1]* ([1]University of Pennsylvania, [2]University of North Carolina at Chapel-Hill) 109

1.60 **Chinese aphasic patients' comprehension deficits with discourse-related constructions.** *Honglei Wang* (Beihang University) 110

1.61 **Distinguishing two routes to silent meaning in the brain.** *E. Matthew Husband[1] and Fernanda Ferreira[2]* ([1]University of Oxford, [2]University of South Carolina) 111

1.62 **Events along the garden path: A reduced N400 and a P600 in semantically reversible discourse.** *Gina Kuperberg and Kristina Fanucci* (MGH/Tufts) 112

2.1   **Discourse-driven biases in native- vs non-native speakers' coreference processing.** *Theres Grüter[1], Hannah Rohde[2], Amy J. Schafer[1]* ([1]University of Hawaii at Manoa, [2]University of Edinburgh)   131

2.2   **Online processing of English garden-path sentences by L2 learners: A visual world study.** *Lucia Pozzan and John Trueswell* (University of Pennsylvania)   132

2.3   **Word order and interference in online gap-filling by bilinguals.** *Irina Sekerina* (College of Staten Island and the Graduate Center, CUNY)   133

2.4   **Syntactic constraints in the processing of wh-movement by L2 learners.** *Adrienne Johnson, Alonso Canales, Rob Fiorentino, Alison Gabriele* (University of Kansas)   134

2.5   **Unfolding an event differently: An ERP study on L1 and L2 processing of grammatical aspect.** *Shengyan Long, Manami Sato, Hiromu Sakai* (Hiroshima University)   135

2.6   **Predictive use of case marking during sentence comprehension: An eye-tracking study of Turkish-speaking children (and adults).** *Duygu Özge[1,2], Aylin Küntay[1], Jesse Snedeker[2]* ([1]Koç University, [2]Harvard University)   136

2.7   **Sarcasm: Do you hear it now?** *Sara A. Peters[1,2], Kathryn Wilson[2], Amit Almor[2]* ([1]Newberry College, [2]University of South Carolina)   137

2.8   **L2 processing of Arabic derivational morphology.** *Suzanne Freynik and Polly O'Rourke* (University of Maryland)   138

2.9   **Referential ambiguity and pronoun resolution: Evidence from pupillometry.** *Manizeh Khan and Jesse Snedeker* (Harvard University)   139

2.10   **The effect of phrase length on the form of referring expressions.** *Hossein Karimi[1], Kumiko Fukumura[2], Martin Pickering[3], Fernanda Ferreira[1]* ([1]University of South Carolina, [2]University of Strathclyde, [3]University of Edinburgh)   140

2.11   **The myth of the Overt Pronoun Constraint in Spanish.** *Carlos Gelormini[1], David Huepe[2], Eduar Herrera[3], Timothy W. Boiteau[4], Margherita Melloni[1], Facundo Manes[1], Adolfo Garcia[5], Agustin Ibañez[1]* ([1]Institute of Cognitive Neurology, Buenos Aires, [2]Universidad Diego Portales, [3]Universidad Autónoma del Caribe, [4]University of South Carolina, [5]Universidad Nacional de Córdoba)   141

2.12   **What types of lexical information are reaccessed during pronoun processing?** *Sol Lago[1], Shayne Sloggett[2], Wing Yee Chow[1], Colin Phillips[1]* ([1]University of Maryland, [2]University of Massachusetts Amherst)   142

2.13   **Disfluency primes.** *Sarah Brown-Schmidt* (University of Illinois)   143

2.14   **Contextual effects on the comprehension of speaker corrections: An ERP study.** *Justine VanDyke-Lyon[1], E. Matthew Husband[2], Fernanda Ferreira[1], Nathan D. Maxfield[3]* ([1]University of South Carolina, [2]St. Hugh's College, Oxford, [3]University of South Florida)   144

2.15   **Lexical disambiguation using parafoveal information.** *Rukshin Shaher and Shravan Vasishth* (University of Potsdam)   145

2.16   **Morphological activation during spoken word recognition in Hebrew.** *Daphna Heller[1] and Avital Deutsch[2]* ([1]University of Toronto, [2]Hebrew University of Jerusalem)   146

2.17   **Predictability and prediction: Are upcoming words pre-activated during sentence processing?** *Wonil Choi and Peter Gordon* (University of North Carolina at Chapel Hill)   147

2.18   **The influence of context information on vocabulary acquisition in reading.** *Randy Lowell and Robin K. Morris* (University of South Carolina)   148

2.19   **Lexical clustering in efficient language design.** *Kyle Mahowald[1], Steven T. Piantadosi[2], Edward Gibson[1]* ([1]MIT, [2]University of Rochester)   149

2.20   **A new account of spillover effects in reading: Evidence from parafoveal masking.** *Michael Shvartsman, Richard Lewis, Satinder Singh* (University of Michigan)   150

2.21   **Auditory confusability vs. phonological neighborhood in language production.** *Susanne Gahl[1] and Julia Strand[2]* ([1]UC Berkeley, [2]Carleton College)   151

2.22   **How modular is lexical category disambiguation?** *Peter Baumann* (Northwestern University)   152

2.23   **Speaker distraction interrupts prosodic cues to discourse status.** *Jennifer E. Arnold, Giulia C. Pancani, Elise C. Rosa* (UNC Chapel Hill)   153

2.24   **Consequences of 'music to one's ears': Structural integration priming from music to language.** *Mythili Menon and Elsi Kaiser* (University of Southern California)   154

2.25   **Rapid adaptation in the pragmatic interpretation of contrastive prosody.** *Chigusa Kurumada[1], Meredith Brown[2], Michael Tanenhaus[2]* ([1]Stanford University, [2]University of Rochester)   155

2.26   **Predicting upcoming words but not semantic features: Evidence from ERPs.** *Nayoung Kwon[1], Pan Liu[2], Patrick Sturt[3]* ([1]Konkuk University, [2]Nanyang Technological University, [3]University of Edinburgh)   156

2.27   **Self or other: Interplay of verb biases and syntactic constraints during reflexive processing.** *Xiao He and Elsi Kaiser* (University of Southern California)   157

2.28   **Expectation adaptation for clustering of syntactic structures.** *Mark Myslín and Roger Levy* (UC San Diego)   158

2.29   **How speakers trade accuracy for speed when producing subject-verb agreement.** *Laurel Brehm and J. Kathryn Bock* (University of Illinois)   159

2.30 **Recent experience changes production preferences in the face of semantic biases.** 160
Victor Ferreira and Liane Wardlow (UC San Diego)

2.31 **Advance planning of verbs in head-final language production.** *Shota Momma, Robert* 161
Slevc, Colin Phillips (University of Maryland)

2.32 **Silent structures in ellipsis: Evidence from syntactic priming.** *Ming Xiang, Julian Grove,* 162
Jason Merchant, Genna Vegh, Stefan Bartell, Katina Vradelis (University of Chicago)

2.33 **Planning units in Tagalog sentence production: Evidence from eye tracking.** *Sebastian* 163
Sauppe[1], Elisabeth Norcliffe[1], Agnieszka E. Konopka[1], Robert D. Van Valin, Jr.[1,2,3], Stephen C.
Levinson[1,4] ([1]Max Planck Institute for Psycholinguistics, [2]Heinrich Heine University, [3]University
at Buffalo, The State University of New York, [4]Radbound University)

2.34 **The upside of not having a syntactic choice: Effects of syntactic flexibility on Korean** 164
production. *Heeju Hwang and Elsi Kaiser* (University of Southern California)

2.35 **Towards the understanding of the correspondence relationship between language-** 165
related ERP components and oscillatory activities. *Hiroaki Oishi[1], Nobuyuki Jincho[1], Reiko*
Mazuka[1,2] ([1]RIKEN Brain Science Institute, [2]Duke University)

2.36 **Are our eyes really faster than our brains? Aligning eye-tracking and ERP time** 166
estimates. *Wing Yee Chow[1], Colin Phillips[1], Suiping Wang[2]* ([1]University of Maryland, [2]South
China Normal University)

2.37 **Eyetracking evidence for the subject relative advantage in Mandarin.** *Lena Jäger[1], Shravan* 167
Vasishth[1], Zhong Chen[2], Chien-Jer Charles Lin[3] ([1]University of Potsdam, [2]Cornell University,
[3]Indiana University)

2.38 **Discourse accessibility and structural bias: Processing D-linked phrases in sluices.** 168
Jesse Harris (Pomona College)

2.39 **New evidence on D-linking.** *Grant Goodall* (University of California, San Diego) 169

2.40 **Hidden factors in the production of grammaticality judgments.** *Gisbert Fanselow[1], Jana* 170
Häussler[1], Thomas Weskott[2] ([1]University of Potsdam, [2]University of Göttingen)

2.41 **The underlying cognitive components of sentence processing: Not all P600s are alike.** 171
Polly O'Rourke (University of Maryland, Center for the Advanced Study of Language)

2.42 **The processing of raising and nominal control.** *Patrick Sturt[1] and Nayoung Kwon[2]* 172
([1]University of Edinburgh, [2]Konkuk University)

2.43 **Biases in resolving wh-dependencies in a hybrid language.** *Dustin Chacón and Colin* 173
Phillips (University of Maryland)

2.44	**Argument-structure driven parsing in Tagalog.** *Michael Frazier and Masaya Yoshida* (Northwestern University)	174
2.45	**Effects of syntactic complexity and animacy on the initiation times for head-final relative clauses.** *Chien-Jer Charles Lin* (Indiana University)	175
2.46	**Collectivity and concreteness in optional Persian number agreement.** *Aazam Feizmohammadpour and Wind Cowles* (University of Florida)	176
2.47	**Number agreement without surface syntax.** *Ming Xiang and Genna Vegh* (University of Chicago)	177
2.48	**Predictability effects of case-marked direct objects: Evidence from Romanian.** *Sofiana Chiriacescu* (University of Köln, Transilvania University of Brasov)	178
2.49	**The locus and nature of the object-extracted relative clause penalty.** *Jeffrey Witzel[1] and Kenneth Forster[2]* ([1]University of Texas, Arlington, [2]University of Arizona)	179
2.50	**Who did what to whom? An investigation of syntactic reanalysis in English and Mandarin.** *Yi Ting Huang[1], Xiangzhi Meng[2], Kathryn Leech[1]* ([1]University of Maryland College Park, [2]Peking University)	180
2.51	**The use of non-structural cues in reflexive resolution: Evidence from eye-tracking.** *Lena Benz[1], Lena Jäger[1], Shravan Vasishth[1], Philip Hofmeister[2]* ([1]University of Potsdam, [2]University of Essex)	181
2.52	**Highs and lows in English attachment.** *Nino Grillo[1], Andrea Santi[2], Bruno Fernandes[1], João Costa[1]* ([1]Universidade Nova de Lisboa, [2]University College London)	182
2.53	**Illusory NPI licensing: Now you see it, now you don't.** *Dan Parker, Glynis MacMillan, Colin Phillips* (University of Maryland)	183
2.54	**Information structure and the 'height' of ellipsis.** *Timothy Dozat[1] and Jeffrey Runner[2]* ([1]Stanford University, [2]University of Rochester)	184
2.55	**Here comes the subject: Listeners use number-marked verbs to predict subject number.** *Cynthia Lukyanenko and Cynthia Fisher* (University of Illinois, Urbana-Champaign)	185
2.56	**An ACT-R model interfacing eye movements with parsing.** *Felix Engelmann, Shravan Vasishth, Ralf Engbert, Reinhold Kliegl* (University of Potsdam)	186
2.57	**Form-based syntactic expectations affect the duration of early fixations in reading.** *Thomas Farmer[1], Klinton Bicknell[2], Michael Tanenhaus[3]* ([1]University of Iowa, [2]UC San Diego, [3]University of Rochester)	187
2.58	**Electrophysiological response to manipulation of syntactic expectations.** *Joe Kirkham, Chelsea Guerra, Edith Kaan* (University of Florida)	188

3.16 **Indefinite NPs introduce new referents but not immediately.** *Maria Luiza Cunha Lima[1], Amit Almor[2], Evgenia Borschevskaya[2], Timothy W. Boiteau[2]* ([1]Universidade Federal de Minas Gerais, [2]University of South Carolina)   226

3.17 **The cost of unexpected contrast: Processing 'let alone'.** *Jesse Harris* (Pomona College)   227

3.18 **Effects of novelty and givenness on acoustic reduction.** *Lap-Ching Keung and Jennifer E. Arnold* (University of North Carolina, Chapel Hill)   228

3.19 **Auditory priming affects planning and execution separately.** *Jason Kahn and Jennifer Arnold* (University of North Carolina, Chapel Hill)   229

3.20 **Form interference effects during silent reading.** *Iya Khelm, Naoko Witzel, Jeffrey Witzel* (University of Texas, Arlington)   230

3.21 **It's probably porridge: The role of tonal probability in Mandarin lexical access.** *Seth Wiener and Kiwako Ito* (The Ohio State University)   231

3.22 **Effects of context and individual differences on processing taboo words within sentences.** *Adina Raizen, Cassie Palmer-Landry, Kiel Christianson* (University of Illinois, Urbana-Champaign)   232

3.23 **Topic, empathy, and point of view.** *Laura Kertz and Corey Cusimano* (Brown University)   233

3.24 **Frequency and distribution of *some* (but not *all*) implicatures.** *Judith Degen, Michael K. Tanenhaus, Christine Gunlogson* (University of Rochester)   234

3.25 **Effects of speaker identity on processing rude and polite language: Evidence from a Twitterish paradigm.** *James Nye, Steven Luke, Justine VanDyke-Lyon, Fernanda Ferreira* (University of South Carolina)   235

3.26 **Facial feedback and the real time comprehension of emotional language.** *Seana Coulson, Joshua Davis, Piotr Winkielman* (University of California, San Diego)   236

3.27 **Shifting viewpoints: Free indirect discourse and sensitivity to perspective-taking.** *Elsi Kaiser, Alexa Cohen, Emily Fedele* (University of Southern California)   237

3.28 **Visuospatial grouping influences expectations about upcoming discourse.** *Elsi Kaiser and David Cheng-Huan Li* (University of Southern California)   238

3.29 **Evidence for a rational probabilistic account of Gricean implicatures.** *Daniel Grodner[1] and Benjamin Russell[2]* ([1]Swarthmore College, [2]Brown University)   239

3.30 **Sensitivity to local discourse vs. global communicative context in gradable adjectives.** *Christina Kim[1], Andrea Beltrama[1], Kristen Syrett[2], Ming Xiang[1], Chris Kennedy[1]* ([1]University of Chicago, [2]Rutgers University)   240

3.31	**Objects and actions in dis-agreement.** *Jason Schoenberg and Heidi Lorimor* (Bucknell University)	241
3.32	**Filling in the blanks in morphological productivity: A word-completion task.** *Kyle Mahowald, Timothy O'Donnell, Joshua Tenenbaum* (MIT)	242
3.33	**How different levels of syntactic flexibility influence language production in Mandarin.** *Xin Zhao and Elsi Kaiser* (University of Southern California)	243
3.34	**Implicit naming in the visual world paradigm.** *Daniel Pontillo, Anne Pier Salverda, Michael Tanenhaus* (University of Rochester)	244
3.35	**Theory of mind drives efficient language production.** *Peter Graff[1], Zoe Snape[1], Jeremy Hartman[2], Edward Gibson[1]* ([1]MIT, [2]U Mass Amherst)	245
3.36	**Individual differences in reading styles and the use of implicit causality as a pronoun resolution cue.** *Arnout Koornneef and Ted Sanders* (Utrecht University)	246
3.37	**Case licensing in processing: Evidence from German.** *Shayne Sloggett* (UMass, Amherst)	247
3.38	**How bizarre: Sentence processing and memory.** *Peter C Gordon, Matthew W. Lowder, Miri Besken, Neil Mulligan* (University of North Carolina, Chapel Hill)	248
3.39	**The parallel computation of phrasal and nonphrasal constituents: Evidence from embedded adjectives in compound nouns.** *Cara Tsang and Craig Chambers* (University of Toronto)	249
3.40	**Individual differences in sentence processing: Separable effects of knowledge and processing skill.** *Peter C. Gordon, Wonil Choi, Renske S. Hoedemaker, Matthew W. Lowder* (University of North Carolina, Chapel Hill)	250
3.41	**On the role of working memory capacity when prediction is not met: Evidence from NPI-processing.** *Juliane Domke* (Humboldt University)	251
3.42	**Working memory and syntactic islands revisited.** *Edward Gibson[1] and Greg Scontras[2]* ([1]MIT, [2]Harvard University)	252
3.43	**How specific should I be? The optimal amount of information in online language comprehension.** *Si On Yoon and Sarah Brown-Schmidt* (University of Illinois, Urbana-Champaign)	253
3.44	**The realization of scalar inferences: Context sensitivity without processing cost.** *Stephen Politzer-Ahles and Robert Fiorentino* (University of Kansas)	254
3.45	**Eye movements reveal causes of delay in negative sentence processing.** *Ye Tian[1], Richard Breheny[1], Heather Ferguson[2]* ([1]University College London, [2]University of Kent)	255

3.46   **Predictive computations underlie the N400's sensitivity to thematic role-reversals.** *Wing Yee Chow[1], Colin Phillips[1], Suiping Wang[2]* ([1]University of Maryland, [2]South China Normal University)   256

3.47   **How hugging differs from giving a hug: Syntax, semantics or mapping.** *Eva Wittenberg and Jesse Snedeker* (Harvard University)   257

3.48   **Semantic effects on anaphor processing.** *Sara Peters[1], Timothy W. Boiteau[2], Amit Almor[2]* ([1]Newberry College, [2]University of South Carolina)   258

3.49   **Advantages of extending vs. mixing metaphors: An ERP study.** *Les Sikos[1], Paul Thibodeau[2], Cassandra Strawser[1], Frank Durgin[1]* ([1]Swarthmore College, [2]Stanford University / Trinity University)   259

3.50   **Lexically predicting visual features of word referents.** *Tristan Davenport, Seana Coulson, Vicky Tu, Benjamin Bergen* (University of California, San Diego)   260

3.51   **Semantic commitment in online verb processing.** *Nicholas Gaylord[1], Micah Goldwater[2], Colin Bannard[1], Katrin Erk[1]* ([1]University of Texas, Austin, [2]Northwestern University)   261

3.52   **Regeneration in verb phrase ellipsis resolution.** *Suzanne Belanger and Ron Smyth* (University of Toronto)   262

3.53   **Two flavors of long distance dependency discerned through island effects.** *Dan Parker and Bradley Larson* (University of Maryland)   263

3.54   **Feedback, risk sensitivity and response-contingent financial payoffs affect reading time for syntactically ambiguous sentences.** *Luis Chacartegui-Quetglas and Colin Bannard* (University of Texas, Austin)   264

3.55   **A rational account of regressions in syntactically complex sentences.** *Klinton Bicknell and Roger Levy* (University of California, San Diego)   265

3.56   **Online filler-gap dependency formation and that-trace effect.** *Morgan Purrier, Masaya Yoshida, Lauren Ackerman, Rebekah Ward* (Northwestern University)   266

3.57   **The role of morphology in phoneme prediction: Evidence from MEG.** *Allyson Ettinger, Tal Linzen, Alec Marantz* (New York University)   267

3.58   **Verb-argument processing with and without event-related knowledge impairment.** *Michael Walsh Dickey and Tessa Warren* (University of Pittsburgh)   268

3.59   **MEG evidence for immediate reference resolution within a visual world.** *Christian Brodbeck and Liina Pylkkänen* (New York University)   269

3.60 **Lexical processing and working memory in individuals with and without aphasia.** *Maria Ivanova[1], Olga Dragoy[2], Svetlana Kuptsova[1], Anastasia Ulicheva[3], Anna Laurinavichyute[4], Lidia Petrova[1]* ([1]Center of Speech Pathology and Neurorehabilitation, [2]Moscow Research Institute of Psychiatry, [3]University of Hong Kong, [4]Higher School of Economics) 270

3.61 **Neural correlates of sentence plausibility in garden-path processing.** *Dirk-Bart Den Ouden, Svetlana Malyutina, Victoria Sharpe* (University of South Carolina) 271

Should psycholinguistics ignore the language of the brain?
Peter Hagoort (Max Planck Institute for Psycholinguistics)
peter.hagoort@donders.ru.nl

From a functionalist perspective all that brain research is claimed to have told us is that language processing "happens somewhere north of the neck" (Jerry Fodor, 1999). I will argue why I disagree with this conclusion, for at least the following three reasons. First, one fundamental question in the language sciences is: what makes the human brain language-ready? Understanding the neural architecture that supports human language function is a crucial part of the explanandum. I will show some unique features of human perisylvian cortex based on data from Diffusion Tensor Imaging and resting state fMRI. The second argument is that even if one is only interested in the cognitive architecture of language comprehension and production, relevant evidence can be obtained from neurobiological data, both structural and functional. I will discuss the consequences of connectivity patterns in the brain for assumptions in processing models of language, and I will show fMRI data based on a repetition suppression paradigm that provide evidence for the claim that syntactic encoding and parsing are based on the same mechanism. Finally, I will argue that framing theories of sentence processing in a way that connects to other areas of cognitive neuroscience might be helpful in asking interesting and relevant new questions. I will illustrate this in the context of the Memory, Unification and Control (MUC) model of language.
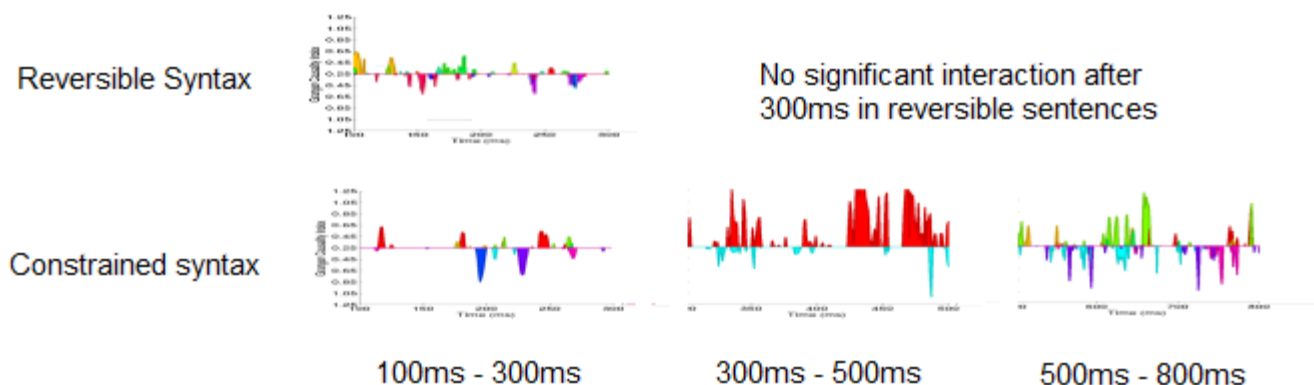
Sentence type effects in Granger causality analysis of MEG/EEG signal

David Caplan (MGH), David Gow (MGH), Reid Vancelette (MGH), Conrad Nied (MGH)
dcaplan@partners.org

There are different effects of semantic constraint in verification and plausibility judgment tasks. In verification, participants must verify the syntactic origin of a plausible meaning in relation to other plausible meanings, requiring a check of reversible sentences. In plausibility judgment, readers must verify the syntactic origin of a plausible meaning against implausible alternatives, or vice versa, requiring a check of constrained sentences. Results of eye fixation studies are consistent with these effects appearing in the main clause of sentences with relative clauses on the matrix subject. Traxler et al (2002) found longer reading times and more regressions from these points in object relative sentences with reversible compared to constrained NPs in verification. In plausibility judgment, Caplan and Evans (2012) found longer total reading time and more regressions from the main verb and longer total reading time for the sentence-final NP in constrained object relative sentences.

We used Granger causality of MRI-constrained EEG/MEG data to examine the hypothesis that listeners test plausibility by comparing emerging syntactic and semantic/thematic analyses during sentence processing. We reasoned that checking of sentences for their syntactic original would involve reciprocal Granger causality between brain areas that support syntactic processing and brain areas that support semantic memory, where the plausibility of a set of thematic roles could be determined. We assumed that the first includes L pars opercularis and triangularis and the second includes L MTG. Greater reciprocal Granger causality between these regions during presentation of the main clause was predicted in constrained than reversible sentences.

Twelve right handed B.U. students made plausibility judgments about spoken semantically reversible and constrained plausible and implausible subject and object relative sentences. EEG/MEG signal was recorded and mapped onto MR images as in previous studies (see Gow and Caplan, 2012, for methods). Granger causality effects were identified in three temporal intervals after the onset of each content word – 100ms – 300ms; 300ms-500ms; 500-800ms – chosen to reflect epochs corresponding to primarily lexical processing, and early and late syntactic processing. Granger causality effects at a threshold of alpha = 0.05 between L pars opercularis/triangularis and L MTG were greater at the main verb and sentence-final noun of constrained than reversible sentences. Granger causality is shown around the axis representing the significance threshold, with positive granger effects towards the L MTG above the axis and positive granger effects from the L MTG to the L IFG shown in reflection below. Colors represent different sub-regions.

Reversible Syntax — No significant interaction after 300ms in reversible sentences

Constrained syntax

100ms - 300ms   300ms - 500ms   500ms - 800ms

The results are the first to document Granger causality effects of incremental sentence processing. They document task effects that need to be considered in the interpretation of other neural and behavioral results.

References

Caplan and Evans, Psychonomics Society, 2012; Gow and Caplan, Frontiers Psychol, 2012, doi: 10.3389/fpsyg.2012.00506; Traxler et al, JML, 47, 69–90, 2002

The P600 indexes rational error correction within a noisy-channel model of human communication

Edward Gibson (MIT), Laura Stearns (Wellesley), Leon Bergen (MIT), Marianna Eddy (US Army), Ev
Fedorenko (MIT)
egibson@mit.edu

When originally discovered, the N400 ERP component was hypothesized to index the ease of lexical access and/or integration of the word's meaning into the preceding context (e.g., *I take my coffee with cream and sugar / **dog***; Kutas & Hillyard, 1980), and the P600 was hypothesized to reflect syntactic integration difficulty (e.g., *Every Monday he mows / **mow** the lawn*; Osterhout & Holcomb, 1992; Hagoort & Brown, 1993). However a number of more recent findings have posed problems for the traditional interpretation of the P600 component (see e.g., Brouwer et al., 2012, for a review). For example, sentences like "*The hearty meal was **devouring**…*" elicit a P600 in spite of being syntactically well-formed (Kim & Osterhout, 2005; Kuperberg et al., 2003). Because these sentences are semantically anomalous, the traditional approach predicts an N400, not a P600. Furthermore, a P600 is generated for spelling errors (e.g., "fone" instead of "phone"; Munte et al., 1998), and no P600 is generated for syntactic errors in "Jabberwocky" sentences (Yamada & Neville, 2007). None of these results can be accounted for by the traditional approach.

We propose and test a novel account of the P600 within a rational inference approach to sentence comprehension (Shannon, 1949; Levy et al., 2009). In a linguistic exchange, the comprehender's optimal guess of what was intended takes into account the probability of the speaker's intended meaning (s_i) given the perceptual input (s_p): $P(s_i \mid s_p)$. By Bayes' rule, this is achieved by multiplying the prior (i.e., what is likely to be said), $P(s_i)$, with the likelihood that a noise process would generate s_p from s_i, $P(s_i \rightarrow s_p)$. A plausible interpretation of the P600 in this communication-based proposal is as an *error correction* process: when a detected error can be corrected, a P600 ensues. If the error cannot be plausibly corrected, then it is possible that an N400 might ensue. This account straightforwardly explains the results from the literature: (1) a P600 occurs for the "traditional" syntactic errors (e.g., number/gender agreement, etc.) because there is a close alternative in these cases, which the listener/reader can infer and correct to (e.g., "mow" can be corrected to "mows" in the example above); (2) an N400 occurs for "traditional" semantic errors because no simple error correction is possible in those cases (e.g., if someone says "I take my coffee with cream and dog", there is no obvious way to arrive at a more plausible target meaning); (3) a P600 occurs for the "semantic P600" examples because there is a close alternative that the producer may have meant (e.g., "*devoured*" for "*devouring*" above; see Kim & Sikos, 2011, for further evidence supporting this proposal); (4) a P600 occurs for spelling errors because it is clear what the intended word is; and (5) a smaller or no P600 is observed for errors in Jabberwocky sentences because it is difficult or impossible to infer plausibly intended meanings because the materials are by design devoid of meaning. We directly evaluated our proposal in a new ERP experiment (29 participants; EEG recording from 32 scalp sites; 160 items; 320 fillers) with 4 conditions, as in (1):

(1) The storyteller could turn any incident into an amusing …

Control **anecdote** / *critical* **antidote** / *syntactic control* = **anecdotes** / *N400 control* = **hearse**

In the critical condition, the target word was semantically implausible but was phonologically and orthographically close to a semantically plausible neighbor (anecdote / antidote in (1)), as established in a norming study. The existence of such a neighbor makes the plausibly intended word recoverable. The noisy-channel account therefore predicts a P600 in this condition.

Results. As expected, a P600 was observed for the syntactic control (anecdotes vs. anecdote; p<0.001), and an N400 – for the semantic control (hearse vs. anecdote; p<0.001). Critically, a P600 was observed in the critical condition (antidote vs. anecdote; p<0.001), similar to that elicited by the typical syntactic violations (number agreement errors in our materials). These results support the rational error correction interpretation of the P600 component.

To a man with a hammer everything looks like a nail: How explicit lexical predictions influence sentence processing – an ERP study

Jakub Szewczyk (Jagiellonian University) & Herbert Schriefers (Radboud University)
jakub.szewczyk@gmail.com

It is now widely established that the language comprehension system actively predicts specific upcoming words, given some constraining context. This effect has been demonstrated for the predicted word itself, or for words closely associated with the predicted word. But this is only one side of the coin. The present study tests whether and how lexical expectations modify the way in which the brain interprets unfolding sentences (often including words unrelated to the predicted word). To this aim, participants were presented with short stories:

During a school trip to the mountains, Peter was bragging about his good sense of direction. Even when he was not able to see the location of the sun, he was always able to point to the north. However, after two days, it was discovered that he was cheating. [prime / no prime]
Somebody noticed that he had hidden a <u>compass</u> *[c] / a* <u>roof</u> *[ic] up in his sleeve.*

All the stories were presented on the screen all-at-once, except for the final sentence, which was presented word-by-word. In half of the stories, just before the story-final sentence (marked by [prime / no prime] above), explicit information was introduced, telling the participants that in the following sentence a specific target word will appear (prime condition). In the other half of items, no information concerning the target word was given (no prime condition). The target word was always the direct object of the main clause of the story-final sentence. In addition to the priming manipulation, congruity of the target word was manipulated: in half of the items, the target word was semantically congruent with the preceding context (congruent condition), while in the other half it was not (incongruent condition). Priming and congruity manipulations were fully crossed.

We focused primarily on the analysis of ERPs elicited by the words of the story-final sentence preceding the target word presentation (hereafter called intervening words), to see if they are processed differently in prime and no prime conditions. The congruity manipulation provided some variation with respect to how related the target word was to the intervening words. We also gathered additional norming data on target word likeliness by truncating the story at 4 positions of the story-final sentence, and asking participants how much the target word would fit, if it occurred somewhere in the remaining part of the story-final sentence. This enabled us to substitute Congruity with the factor Likeliness of the target word at each tested position, based on trichotomized values of these ratings.

Until the presentation of the main verb, priming led to a sustained negativity, with a fronto-central distribution at the first intervening word that soon changed into a parietal distribution at subsequent intervening words. In the primed condition, at the first intervening word there was no effect of target word Likeliness (measured at this position), but it emerged at least at the word directly preceding the main verb (with a central distribution), leading to more negative ERPs for positions at which target word was not rated as likely to occur in the following part of the sentence. The effect of Likeliness was the most extreme at the verb, leading to a full-blown N400 for target nouns rated to be unlikely, and to a full reduction of the negativity (to the level of ERPs in the unprimed condition) for target nouns rated as likely to occur next. At the target word this pattern broke down: unlikely unprimed target words led to a standard N400 followed by a P600, whereas primed words rated as unlikely led to no N400 and no P600. Primed and unprimed likely target words led to similar ERPs.

These results suggest that: 1) before hitting the target noun, prediction came with a cost, particularly when the predicted element was not very likely (the only situation when there was no cost associated with prediction at the intervening words, was at the verb, when the target word was rated as likely), 2) the prediction cost for unlikely target words turns into a benefit when the target word is presented, leading to a full reduction of the N400, 3) the only reason why unlikely nouns led to the N400 at the (fully congruent) verb could be that it resulted from a difficulty of integrating the target noun with the context, 4) the negativity occurring before the verb resulting from priming was not the standard N400 effect, because of a different scalp distribution.

Exploring representations of event duration in language
Gitte Joergensen & Silvia Gennari (University of York)
g.joergensen@psych.york.ac.uk

Previous studies have shown that durative events (to owe money) take longer to read than punctual events (to lose money) (Coll-Florit & Gennari, 2011). Similarly, in narrative comprehension, larger temporal distances between events take longer to process (Zwaan, 1996; Kelter, Kaup and Claus, 2004). This suggests that longer events and temporal distances recruit more semantic information during processing. However, distance effects in narratives may be due to recruiting knowledge of causal connections between events (the longer the connection, the longer the processing). Similarly, comparisons across different verbs (to owe vs. to lose money) may be due to differences in causal event structure. Therefore, it remains unclear what type of information is recruited when processing long events.

To address this issue, we constructed materials like (1) in which the discourse relations and the event referred to stay the same and only the event duration interpretation varies due to minimal changes in the preceding context (plausibility didn't differ).

1. Lisa was moving to a new flat near the university.

    Long/short condition: John spent his morning/an hour there.

    He spent all that time assembling her bed.

In Experiment 1, we tracked participants' eye-movements while looking at objects on the screen containing only one object related to the story being heard (e.g. bed in (1)). Participants only heard one condition for each item (long or short). Results indicated that first fixation durations on the relevant object (bed) were longer for the long condition while hearing "her bed" ($p < .05$). These results were replicated (Experiment 2) with a different set of materials in which the scale of the events' duration was longer (e.g., building a house in two weeks vs. a month).

In Experiment 3, we used the same stimulus materials in a probe recognition task. After reading stories like (1) (with an additional final sentence to avoid recency), participants were presented with words (e.g. bed), and were instructed to indicate whether the word had occurred in the story. The word probes could either be from the critical verb phrase ("assembling her bed") –late probes– or from the beginning of the story ("flat") –early probes–, in which case no effect of duration should be observed. This manipulation thus serves to address alternative explanations of Experiments 1 and 2, in that the results should not be driven by the mere presence of an adverbial temporal phrase. We found as expected, that participants were faster in recognizing late probes in the short-version of the story than in the long-version ($p < .05$), but no difference was found for early probes. This indicates that longer events are less accessible from memory.

These results indicate that event duration effects prevail even when the same verb and narrative structure are used in the stimuli. This suggests that the representation of an event's internal development is more complex for longer events. We argue that understanding longer events recruits experience-based knowledge of the sub-events that would likely occur, given the context, thus leading to more processing cost.

References

Coll-Florit, M. and Gennari, S. P. (2011) Time in language: event duration in language comprehension, Cognitive Psychology, 62, 41-79.
Zwaan, R. A. (1996) Processing narrative times shifts. Journal of Experimental Psychology: LMC, 22, 1196-1207.
Kelter, S., Kaup, B., & Claus, B. (2004) Representing a described sequence of events: A dynamic view of narrative comprehension. Journal of Experimental Psychology: LMC, 30(2), 451-464.

Bi-directional structural priming between mathematics and language
Christoph Scheepers (University of Glasgow) & Patrick Sturt (University of Edinburgh)
patrick.sturt@ed.ac.uk

Scheepers et al. (2011) showed that the structure of a correctly solved mathematical equation affects how people complete a subsequent sentence containing high vs. low relative-clause attachment ambiguities. Here we investigated whether such cross-domain structural priming effects generalize to different structures and tasks, and importantly, whether they also hold in the reverse direction.

In two questionnaire studies, mathematical equations like (1a) versus (1b) were crossed with pre-tested noun-phrases like (2a) (encouraging a right-branching interpretation: [*divorced* [*hospital nurse*]]), versus (2b) (encouraging a left-branching interpretation: [[*dental hospital*] *nurse*]). Note that (1a) is structurally congruent with (2a), and (1b) with (2b), respectively (the branching structure of 1a,b is dictated by the relative precedence of the "*" and "+" operators).

1a. $5 + 2 * 7 =$

1b. $5 * 2 + 7 =$

2a. *divorced hospital nurse*

2b. *dental hospital nurse*

In Experiment 1, the equations (1) were used as primes and the noun-phrases (2) as targets, and vice versa for Experiment 2. In both experiments, participants had to solve the equations and provide 5-point sensicality ratings for the noun-phrases. Various filler items were employed to conceal the experimental aims.

Experiment 1 (math -> language priming) tested 36 mathematically 'skilled' participants (indicated via pre-assessment) who solved 97% of the prime-equations correctly. This experiment showed an average increase in sensicality ratings for (2) when the noun-phrases followed structurally congruent prime equations, relative to when they followed incongruent prime equations (4.13 vs. 3.85; ps < .01 by ordinal-logistic *GEEs*). Experiment 2 (language -> math priming) tested 36 mathematically 'less skilled' participants (again, indicated by a pre-assessment) and showed that the likelihood of correctly solving the equations in (1) was higher when they followed structurally congruent than incongruent prime noun-phrases (65% vs. 53%; ps < .01 by binary-logistic *GEE*s).   This effect remained reliable in a subsequent analysis including only structural errors (i.e. errors resulting from calculating the equation using the wrong branching structure), and excluding other errors.

These results extend the conclusions from Scheepers et al. (2011) in important ways. The observed bi-directionality of cross-domain structural priming strongly supports the notion of shared syntactic representations (or recursive procedures to generate and parse them) between arithmetic and language, and is compatible with recent brain imaging work suggesting that some aspects of structure building in the two domains may recruit the same cortical regions (see, e.g., Makuuchi et al, 2012). Moreover, the results provide further evidence that the mental representation of syntactic structure can be highly abstract, without being tied to lexical or semantic content.

References

Scheepers, C., Sturt, P., Martin, C.J., Myachykov, A., Teevan, K., and Viskupova, I. (2011). *Psychological Science, 22 (10).* pp. 1319-1326.
Makuuchi, M., Bahlmann, J., & Friederici, A. D. (2012). *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 367(1598),* 2033-2045.

Building meanings in theory vs. in the brain
Liina Pylkkänen (New York University)
liina.pylkkanen@nyu.edu

In order to estimate whether brain evidence supports or contradicts a cognitive model (based on non-neural data), one needs to engage in experiments designed to test the predictions of that model. The degree to which brain research on language has been rooted in theoretical models of the grammar varies between different subfields/representational levels, but this connection has arguably been the weakest for the study of the combinatory semantic system. Little brain research has aimed to connect with models of semantic composition in formal semantics. These models have, however, been the basis of my group's MEG research on semantic composition for the past decade or so and our results have put us in a position to at least discuss the degree to which our findings conform to or challenge various theoretical hypotheses about the architecture of the semantic combinatory system. The critical question is to what extent semantic composition is computationally monolithic (i.e., achieved via a single rule such as function application) or subdivided into different subroutines (achieved via a set of distinct rules). The two types of models are typically considered as opposing views in formal semantics, but our results so far suggest that both may in some sense be true: certain nodes of the network that appears to achieve semantic composition have a computationally highly general profile, whereas other nodes appear specialized for more specific operations. In sum, the overall picture that is emerging is one where composition is achieved by a network of regions which vary in their computational specificity and/or domain generality.

Teasing apart coercion and surprisal: Evidence from ERPs and eye-movements
Francesca Delogu, Heiner Drenhaus & Matthew Crocker (Saarland University)
delogu@coli.uni-saarland.de

Complement coercion refers to expressions in which a verb that semantically selects for an event complement appears with an NP object denoting an entity (e.g., *began the book*). Such expressions have been hypothesized to require type-shifting of the complement noun from an entity to an event or activity (e.g. *began reading the book*) (Pustejovsky, 1995). This view is supported by a number of reading time studies (e.g., Traxler et al., 2002) showing that coercion expressions are harder to process than control expressions that instantiate the default event-sense associated with coercing expressions (e.g., *read the book*). In two subsequent ERP studies (Baggio et al., 2010; Kuperberg et al., 2010), coercion expressions were compared to control expressions (*wrote*) and implausible animacy-violated sentences (*astonished*), as in *'The journalist began/wrote/astonished the article'*. Both studies reported an N400 effect on coerced nouns, which was similar in magnitude and topography to the N400 effect evoked by animacy-violated complement nouns. Based on this finding, Kuperberg et al. (2010) interpreted the N400 modulation to coerced nouns as reflecting the mismatch between the semantic properties of the verb and those of the complement rather than the effort of type-shifting.

In our studies we investigated whether such effects may rather be explained in terms of surprisal (Hale, 2001; Levy, 2008). The predictability of the target noun in control contexts is generally higher than it is in coercion contexts (and in animacy-violation contexts). As a consequence, surprisal for the complement noun – estimated as the inverse of its log probability – will be lower in control contexts than in coercion contexts, suggesting that coercion effects may at least partially be explainable in terms of surprisal. To test this hypothesis, we conducted an ERP and an eye-tracking study in German, testing sentences shown in (1), where coerced nouns were contrasted to a preferred condition instantiating the default interpretation of coercion expressions (*read*) and a 'neutral' non-coercing condition (*bought*) where the verb was equally unconstraining as the coercing verb (thus resulting in similar surprisal).

The standard coercion account (Pustejovsky, 1995) predicts higher processing costs (reflected as an increased N400 or increased reading time) for coerced nouns relative to both preferred and neutral nouns. The surprisal account, in contrast, predicts higher processing costs for coerced and neutral nouns relative to the preferred (and more predictive) nouns, and no differences between coerced and neutral nouns. Consistent with the surprisal account, the ERP study showed a similarly enhanced and distributed N400 for coerced and neutral nouns relative to preferred nouns. The eye-tracking study, however, showed evidence for both surprisal and standard coercion accounts. The first reliable evidence of processing costs emerged in the spillover region (*on holidays*) as a function of surprisal: the analyses revealed more first-pass regressions and longer regression-path times for the coercion (*began*) and the neutral (*bought*) conditions relative to the preferred condition (*read*), and no differences between the coercion and the neutral conditions. Evidence for coercion effects emerged only in later measures of the object region, with longer total reading times for the coercion condition relative to both the neutral and the preferred conditions.

Overall, our results suggest that the coercion cost is initially driven by surprisal for the object noun (as shown by the N400 effect and the early effect on the spillover region) and only in subsequent stages of processing (reflected by total reading times on the object region) are type-shifting operations initiated.

(1) a. John **began** the <u>book</u> on holidays.   (coercion)
   b. John **read** the <u>book</u> on holidays.   (preferred)
   c. John **bought** the <u>book</u> on holidays.   (neutral)

The experiments that we finished: Structural separation reduces the cost of coercion

Matthew W. Lowder & Peter C. Gordon (University of North Carolina at Chapel Hill)
matt.lowder@unc.edu

Verbs like *begin* or *finish* semantically require a complement that represents an action or event, as in *The secretary began writing the memo*. When the complement of these verbs is instead a noun phrase (NP) representing an entity (e.g., *The secretary began the memo.*), the relationship between the verb and complement constitutes a semantic-type mismatch, and the reader must engage in an interpretive process of coercion, where additional semantic material is built into the representation to arrive at the correct interpretation. Several experiments using a variety of methodologies have demonstrated that complement coercion imposes an online processing cost[1-5]. The current pair of eye-tracking experiments investigated the hypothesis that the coercion cost would be reduced when the structure of the sentence deemphasizes the relationship between the verb and the complement NP.

Both experiments involved modifications of a set of stimuli that has previously been shown to elicit coercion effects in eye-tracking[2]. Experiment 1 compared sentences like those in (1), where the verb and target NP were embedded together in a subject-extracted relative clause (SRC; 1a & 1b) or the target NP appeared as the sentence subject and the verb was embedded in an object-extracted relative clause (ORC; 1c & 1d). There was a significant interaction in regression-path duration on the matrix verb, second-pass duration on the target NP, and total time on the target NP. In each of these analyses, the coercion effect was stronger in SRCs than ORCs. Moreover, regression-path duration on the matrix verb showed that ORCs were more difficult than SRCs in the control verb condition (1c vs. 1a); however, there was no difference in the processing of ORCs vs. SRCs in the coercion condition (1d vs. 1b).

Experiment 2 compared sentences like those in (2). This manipulation allowed us to examine the magnitude of the coercion effect when the verb and NP appear together in the same clause (2b vs. 2a) compared to when these two elements appear in separate clauses (2d vs. 2c), but without the differences in word position that existed in Experiment 1. Our results revealed a significant interaction in regression-path duration on a region immediately following the target NP, as well as significant interactions on the target NP itself in both second-pass duration and total time. The source of all interactions was a robust coercion effect when the verb and NP appeared in the same clause (2b vs. 2a) and a coercion effect that was reduced in magnitude or eliminated altogether when the two words were separated by a clause boundary (2d vs. 2c).

We propose that sentence structure acts as a powerful cue to readers, indicating where to focus attention. When a coercion verb and complement NP appear in the same clause, their atypical semantic relationship is particularly salient, which leads to processing difficulty. In contrast, when the two constituents appear in separate clauses, their relationship is deemphasized and is processed more shallowly. This perspective is consistent with various proposals that sentence structure is an important factor that influences the depth at which sentential relations are processed[6-9].

Experiment 1
1a. The secretary that wrote the memo was…
1b. The secretary that began the memo was…
1c. The memo that the secretary wrote was…
1d. The memo that the secretary began was…

Experiment 2
2a. It was the secretary that wrote the memo…
2b. It was the secretary that began the memo…
2c. What the secretary wrote was the memo…
2d. What the secretary began was the memo…

References

[1] McElree, Traxler, Pickering, Seely, & Jackendoff (2001), *Cognition*. [2] Traxler, Pickering, & McElree (2002), *JML*. [3] Pylkkänen & McElree (2007), *JoCN*. [4] Kuperberg, Choi, Cohn, Paczynski, & Jackendoff (2010), *JoCN*. [5] Husband, Kelly, & Zhu (2011), *JoCN*. [6] Ferreira & Patson (2007), *Lang & Ling Comp*. [7] Sanford & Sturt (2002), *TiCS*. [8] Lowder & Gordon (2012), *JML*. [9] Lowder & Gordon (in press), *JEP:LMC*.

Going the distance: Pronoun resolution relies on direct-access retrieval from memory
Stephani Foraker, & Bryan Wight (SUNY Buffalo State)
forakesm@buffalostate.edu

Pronoun resolution relies on working memory operations, as readers need to access referent representations to establish and resolve co-reference. We present two experiments that contrasted three accounts of pronoun resolution. The first is backward-serial search, which posits that readers mentally scan from the pronoun backwards through discourse representations in memory until the referent is reached (in the spirit of Sternberg, 1966). The second is direct-access retrieval, proposing that readers use content-addressable cues at the pronoun (i.e., gender, number, animacy) to directly retrieve the appropriate referent, obviating a search through intervening information. In addition, direct-access retrieval predicts that additional entities with partially-overlapping features should create greater interference with the intended referent, increasing the difficulty of co-reference. Direct-access retrieval underlies comprehension of verb argument dependencies (McElree, Foraker, & Dyer, 2003) and verb-phrase ellipsis (Martin & McElree, 2008, 2009) and has been suggested for inter-sentential dependencies such as co-reference (Foraker & McElree, 2007) but not demonstrated. The third account, Centering theory, states that a pronoun co-refers with the highest-ranked entity in the reader's mental discourse model, from an ordered set of potential referents (Gordon, Grosz, & Gilliom, 1993; Gordon & Scearce, 1995). Several factors contribute to higher ranking, including subject position, used here.

In Experiment 1, 34 participants' eye-movements were tracked as they read 30 discourses (among 90 distracters), normed for equal plausibility, and counterbalanced across lists. Example: *Barbara decorated the house with lights for Christmas short [in December with Steve]long. She hummed contentedly*. The long condition includes more intervening entities, but the referent (*Barbara*) remains the highest ranked entity in both conditions. On the pronoun, following verb, and adverb, we found no differences between conditions for gaze duration, regression path duration, or total time. These findings contradict backward-serial search, but are consistent with either direct-access or Centering theory. Total reading time on the referent, however, was significantly greater in the long condition than short, suggesting resolution of retrieval interference.

As a more stringent test, in Experiment 2, we used the multiple-response speed-accuracy tradeoff procedure to model separate estimates of processing accuracy and speed over time (e.g., McElree et al., 1993). Backward-serial search predicts a slower speed (rate or intercept parameters) for the long condition, as readers would take longer to complete mental scanning. In contrast, direct-access retrieval predicts no difference in the speed parameters, since pronoun cues match the intended referent in both conditions, producing equal speed of referent access. However, the long condition should show lower accuracy (asymptote parameter), due to interference from more entities. Centering theory predicts no differences in speed or accuracy, since the referent is ranked the most prominent entity in both conditions. Fifteen participants read each discourse in RSVP fashion, and entered a series of acceptability judgments for both acceptable (above) and unacceptable continuations (*…She screeched contentedly.*), from which d' accuracy was calculated. The 30 discourses appeared among 90 distracters in each of the counterbalanced sessions. We found support for the direct-access account: the best-fit model had two asymptotes (short = 2.87 d' units, long = 3.07), one rate and one intercept, with a higher adjusted-R^2 (.9947) than the 1-asymptote model (.9906), countering Centering theory predictions. Fits with two rates or two intercepts (.9945) produced inconsistent ordering of the speed parameters (long vs. short, ps > .21), countering serial search predictions.

Taken together, the experiments indicate that pronoun resolution relies on direct-access retrieval of the referent representation, providing further support for the generalizability of this cognitive mechanism. In addition to considering the prominence of the referent, other discourse entities affect co-reference resolution.

Individual differences in anaphoric processing: Insights from mouse-tracking

Elsi Kaiser & Alexis Harper (University of Southern California)
emkaiser@usc.edu

Existing research on reflexives has led to mixed results regarding the role of Binding-theoretic constraints (Sturt 2003 vs. Badecker & Straub 2002). We tested whether individual differences in processing style influence reflexive processing (cf.Yu'10). The idea of different processing styles is central in autism research. The Weak Central Coherence account claims that autistic individuals focus on bottom-up details (e.g., Happe 1999), while "normals" focus more on higher-level/top-down information (e.g., Jolliffe & Baron-Cohen 1999). Importantly, so-called 'autistic traits' are present in the general population, at lower levels, measurable by the Autism Spectrum Quotient questionnaire (AQ, Baron-Cohen et al. 2001).

In the processing of reflexives, are higher AQ-scores correlated with a stronger adherence to bottom-up syntactic cues, and less use of top-down/contextual/real-world knowledge (e.g. verb-directedness)? To test this we used sentences like ex.(1) and manipulated (i) whether the embedded verb was self-directed (e.g., *shave, wash*), other-directed (e.g., *kick, tickle*), or nonsense (no semantic biases), and (ii) whether the matrix subject (not licensed by Binding Theory) matched reflexives' gender. We had 24 targets, 40 fillers and 18 participants (normal 'neurotypical' undergraduates).

Fig 1a

Fig 1b
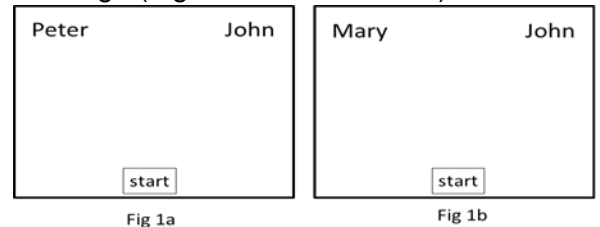
(1) {Mary/Peter} said that John {tickled/shaved/dribbed} himself.

Using mouse-tracking, we recorded mouse-coordinates in real-time (Freeman & Ambady '10). Participants saw words in left and right corners of the screen (matrix and embedded subjects, L/R-balanced), clicked 'start' to hear the sentence, and clicked on the last-mentioned entity (in targets=>whoever 'himself/herself' refers to). Practice trials confirmed people understood task/instructions. Afterwards, participants completed questionnaires, including the AQ. **Analysis.** We measured deflection towards the matrix subject, using the area under the curve, relative to a straight line-to-target (Freeman & Ambady 2010, Freeman's figure below).

Predictions. If anaphoric processing is influenced by processing style, high AQ scorers (more 'autistic traits') are expected to focus mostly on the reflexive (its syntactic requirements) and less on contextual/verb cues. Lower AQ scorers (less 'autistic traits') might be more sensitive to verb information. Specifically: SELF-DIRECTED VERBS: Both low and high AQ-scorers should focus on local subjects (due to verb semantics and reflexive's requirements, respectively). OTHER-DIRECTED VERBS: Low AQ-scorers scorers might be 'tempted' by the matrix subject (verb semantics). High AQ-scorers should focus on the local subject (bottom-up bias). NONCE VERBS: Low AQ-scorers might again consider the matrix subject more than high AQ-scorers.
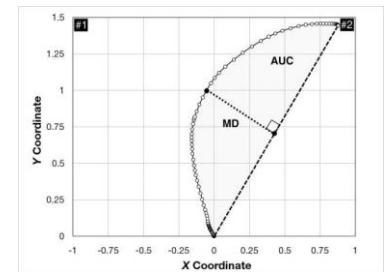
Results. Mouse-tracking reveals significant ($p<.05$) deflection/attraction towards matrix subjects that match the reflexive's gender (for all verb types)--suggesting that processing is not fully constrained by Binding Theory. Furthermore, we find intriguing AQ-related effects. **(i) Self-directed verbs** show no correlation between AQ-scores and deflection/attraction to matrix subject. **(ii) Other-directed verbs** show a negative correlation between AQ and deflection amount (gender-matching subject: $p<.05$, non-gender-matching subject: $p=.064$). **(iii) Nonsense verbs** show a significant negative correlation between AQ and deflection amount (p's$<.03$): The higher the score (more 'autistic traits'), the smaller the amount of deflection. Higher AQ-scorers perform better: They are less 'tempted' by the matrix subject and better able to identify the syntactically-licensed antecedent. In sum, while many questions remain open, this study provides initial, novel evidence of individual cognitive processing differences influencing reference resolution.

Contributions of declarative memory to on-line reference resolution: Findings from amnesia
Sarah Brown-Schmidt (University of Illinois), Jake Kurczek & Melissa Duff (University of Iowa)
sarahbrownschmidt@gmail.com

A growing body of research suggests the hippocampus contributes to a variety of cognitive domains beyond its traditional role in declarative memory. We propose the hippocampus, in its capacity for relational binding, representational flexibility, and on-line maintenance and integration of rich, multimodal relational representations, is a key contributor to language processing and use (Duff & Brown-Schmidt, 2012). Here we test the hypothesis that referential processing, which requires maintaining and integrating representations of potential discourse referents with the unfolding linguistic signal, is hippocampus-dependent. We test for hippocampal involvement in even very short discourses.

Method. Combining eye-tracking and neuropsychological methods, eye-tracked participants (12 young adult control participants, 4 amnesic patients with bilateral hippocampal damage and 8 matched healthy comparison participants) viewed a scene while listening to short dialogs introducing two familiar and highly gender-marked characters and then referred to one using a pronoun; e.g., "*Mickey is racing toward Donald/Minnie, while a house is burning down. He/She is wearing a fire hat, and it looks like they're going to fight the fire*". The design was modeled on previous research with young adults (Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000), and manipulated the gender of the two characters (same vs. different), and the order in which the pronoun's referent was introduced in the story (first vs. second). A minimum of 3 words followed the pronoun before it was fully disambiguated given the scene (e.g., at *fire hat*). Gender of the critical referent was counterbalanced.

Results. Eye-movements immediately following pronoun onset (200-1000ms) were analyzed using mixed-effects models (maximal random effects). Replicating previous work, young adults had two main effects and a gender*order interaction (t=4.42): When the two characters were of the same gender, there was a clear order effect (t=4.08) such that listeners preferred to fixate the first-mentioned referent (e.g., Mickey). In the different-gender condition, there were overall more target fixations, and the order effect was not significant (t=1.31).

Despite the brief nature of the dialogs, these processes drew on declarative memory: Amnesic patients and healthy comparison participants showed significant differences in the magnitude of the order effect (order*participant group t=2.13), and the interaction with gender (gender*group t=2.35). Whereas comparisons performed like young adults (ts>4.5), amnesic patients showed neither an effect of order (t=.62), nor a gender*order interaction (t=1.01), and no significant preference for the 1st mentioned referent in the same-gender condition (t=.87), unlike comparisons (t=4.71). At a later time-region (1000-1800ms), a numeric trend for amnesics to fixate the 1st-mentioned referent in the same gender condition was still not significant (t=1.58).

Conclusion. These findings suggest that amnesic patients experienced difficulty in maintaining and integrating information even over a very short discourse history and with an immediately available visual context. These findings are striking given the traditional view of hippocampus contributing exclusively to long-term memory and of referential processing as relying on the frontal lobes and its putative functions (e.g., working memory). Linking deficits in language processing to the hippocampus demonstrates how promiscuously the hallmark processing features of the hippocampus are used in service of a variety of cognitive domains.

Our findings are consistent with single-store accounts of memory that propose separability between a single item in the focus of attention, and memory, but not a distinction between working and long-term memory (Oztekin, Davachi, & McElree, 2010; also see McElree, 2006; Lewis & Vasishth, 2005). The fact that use of the very recent discourse history requires access to hippocampal representations is consistent with fMRI evidence of hippocampal activity in retrieval of information from memory for all items in a studied list but the most recent (Oztekin, et al., 2010; Oztekin, McElree, Staresina, & Davachi, 2008). Computing the temporal order of the two mentioned items may be a key contributor to the patients' difficulty at resolving competition between the target and competitor, as other findings with this patient population show *success* at distinguishing two identical images when only one had previously been mentioned (Rubin, Brown-Schmidt, Duff, Tranel, & Cohen, 2011).

When timing is (almost) everything: Referential dynamics in parent-child interactions

John Trueswell, Yi Lin (U. Penn.), Erica Cartmill (U.Chicago), Benjamin Armstrong (U. Penn.), Susan Goldin-Meadow (U. Chicago) & Lila Gleitman (U. Penn.)
trueswel@psych.upenn.edu

When talking about our surroundings, extra-linguistic cues to a speaker's referential intent (e.g., gesture, looking at objects) are often quite brief and believed to be precisely timed with linguistic input. Yet, evidence for this precise timing and listeners' sensitivity to this timing has not been studied in any detail, especially in naturally occurring speech outside the lab. This gap in our understanding is surprising when one considers the problem of early word learning: at least for some initial set of words, the only evidence about the meaning of a word comes from this 'buzzing, blooming' situational context, filled with multiple objects, events, attributes, and human behaviors. From this dynamic scene, the infant would need to read referential intent. Laboratory experiments[1] show that infants can indeed read referential intent from head and body posture, but these studies are limited to discriminating between two objects as referents, coupled with a fairly constant physical stance (staring at the object) as the cue. What do the dynamics of reference look like in everyday parent-child interactions? And to what extent are these cues and their timing used even by sophisticated adult observers?

Study 1. We report second-by-second annotation of 40-second videos (N=560) of parent-child interactions in the home, in which parents (N=56) spontaneously uttered one of 41 common concrete nouns (e.g., "ball") to their 14-18 month offspring. We first assessed how well each interaction conveyed referential intent by having naïve adults (N=218) view muted versions of these videos in which a beep was heard at the moment the parent uttered the target word. Observers' accuracy in guessing the word served as a proxy for referential informativity. Replicating past findings[2], highly informative (HI) videos (>50% correct) were rare (only 18% of videos) – attesting to the difficulty of reading referential intentions from natural interactions. Crucially though, we show in time-course plots that HI interactions are characterized by the precise timing of several referential cues, as annotated by two trained (well agreeing) coders. HI interactions had: (1) increased likelihood of the referent being present, going from 74% to 92% presence between -4 & 0 seconds before word onset; (2) increased parent attention to referent, sharply rising during this same period (16% to 58%); (3) increased child attention to referent, starting 8 seconds before onset (20% to 54%); (4) increased parent gesture / manipulation of referent starting 3 seconds before onset (4% to 22%). Reliability of referential informativity on these cues was supported by GLM analyses (all p's<0.0003). Note similar attentional / gestural dynamics occur even for those interactions in which the referent was continuously present throughout the video, revealing contributions of these attentional / gestural cues over and above any simple cue pertaining to the sudden appearance of the referent just before word onset.

Study 2. Here we show that the precise timing of these cues matters for observers to infer referential intent. An additional 48 naïve adults viewed 27 HI videos with the audio muted and participants guessed the word uttered at the beep. Here however, the beep was surreptitiously moved -4, -2, 0, +2 or +4 sec. from the actual word onset. Disruption of the timing of cues relative to word onset dramatically decreased accuracy: the 0 sec offset was 62% correct, as compared to 36%, 43%, 54% and 44% for -4, -2, +2, +4 offset respectively. With the exception of +2 offset, each was significantly different from 0 sec offset condition (p's<0.01).

Summary. Previous work on the reading of referential intent has primarily used artificial laboratory settings and not included detailed analyses of timing. But, for theoretical coherence, one must specify when word and object can be claimed to have occurred "together" in everyday parent-child interaction to support learning. Here we show that high accuracy in referent identification is associated not only with the presence of relevant cues, but also their precise (within 2 sec) alignment with word onset so as to forge the perception of a causal (intentional) link between the two. Children likely monitor for precise relationships between events of the world and word; and, when spotted, these 'epiphany' moments push learning forward.

1. Baldwin, D. A. (1991). *Child Development, 62(5),* 875-890.
2. Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). *Cognition, 73*, 135-176.

Processing of novel compounds in adults and children: One word or two?

Yuki Hirose (The University of Tokyo), Takefumi Ohki (The University of Tokyo), & Reiko Mazuka (RIKEN)
hirose@boz.c.u-tokyo.ac.jp

Recent psycholinguistic studies on compound processing support models where compounds are processed not only as a whole lexical units but also as decomposed forms (Schreuder & Baayen, 1995; Baayen et al., 1997; Baayen, 1992, among others). Studies also show listeners are sensitive to particular phonological patterns characterizing compounds, such as Compound Stress Rules (MaCauley, Hestvik & Vogel 2012 (English), Tsiamas & Kehayia 2009 (Greek)). However, those rules require prosodic information from the second constituent to determine the compound's well-formedness. The Compound Accent Rule (CAR) in Japanese provides a case in which the prosody on the first constituent (C1) alone can provide a cue as to whether it is a part of a compound or a single word before receiving the second constituent (C2), enabling investigation of within-word incremental pre-head processing of compounds. In Japanese, single-word accentuation is lexically determined, whereas for compounds, CAR assigns an accent to the antepenultimate mora, or the original position of C2, and eliminates the original accent on C1. Thus, the surface accent pattern on C1+C2 in (2) remains the same when the C1 is originally unaccented, whereas the original accent on C1 disappears in (1) when it becomes part of a compound (" ' " indicates the position of the lexical accent).

(1) Change in C1 accent :   (e.g.,) *to'mato* (tomato)+ *pa'nda* (panda)→ *tomato pa'nda*
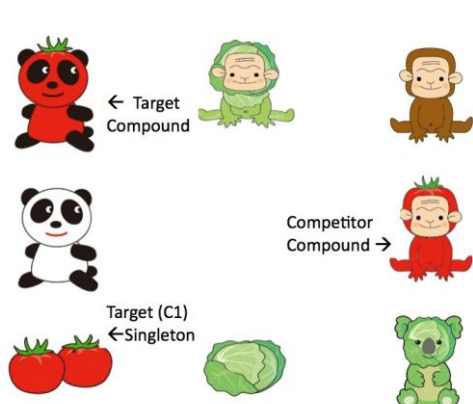(2) No change in C1 accent: (e.g.,) *ringo* (apple)   + *ko'ara* (koala) → *ringo ko'ara*

Two visual world studies on sixteen adults and sixteen children (age: 6-7) using twelve item pairs such as (1) and (2) investigated whether (i) CAR allows immediate by-passing of the decomposed processing of C1, resulting in faster access to the compound meaning, and (ii) whether CAR is used the same way among adults and 6-7 years olds, who both have acquired CAR (Shirose & Kiritani 2001, also tested in our production study). In the experiments, participants saw eight visual objects on the screen, which depicted single animals, fruits, and vegetables such as *to'mato* or imaginary "combined creatures" such as *tomato pa'nda* were presented (as in Figure 1). At the same time, they heard audio stimuli such as "Where is *to'mato*? / *tomato pa'nda*?" in Japanese and were asked to point at the correct referent (Note: Japanese doesn't have number distinction). Of our interest were looks to a) the Target Compound, b) the Target Singleton, and c) a Competitor Compound (sharing the same C1 with the Target Compound) (See Fig.1). We analyzed the log ratio between the looks to the Target Compound object and those to the Target

Figure 1. Sample visual scene for "Where is *to'mato*? / Where is *tomato pa'nda*?"

Singleton, and also the log odds of the looks to the Competitor Compound object against all the other objects combined. The effect of the C1 accent change (e.g., (1) vs. (2)) on these measures were examined over successive 200ms intervals from 300ms after the C1 onset.

For adults, there were more looks to the Competitor Compound for (1) compared with (2) (in a corresponding picture for (2)) (p <0.01), in the 700-900ms interval following the C1 onset. Since C2 of the compound reference should eliminate the Competitor Compound as a possible referent (e.g., "monkey" as in "tomato monkey" in Figure 1), we assume that the effect in this time interval reflects the processing of C1 ("tomato") before the information from C2 is processed, although the C2 has already been heard at this time. In the same window, there were more looks to the Target Compound for (1) compared with (2) (p <0.05). Looks to the Target Compound and of the Target Singleton showed divergence about 200ms faster in (1) compared to (2), presumably due to faster facilitation of compound meaning due to the C1 accent cue. In the same interval (700-900ms following the C1 onset), there were more looks to Target Singleton for (2) compared to (1) (p <0.05), indicating more activation of the single word interpretation of C1. In contrast, children exhibited no comparable effects. Instead, in the post-sentential region (1500-1700ms interval after the C1 onset and onward), there were consistently fewer looks to Target Compound for (1) than for (2) (p <0.05), indicating that accent cue of CAR actually hindered compound interpretation in children.

The results from adults are best explained by a dual route model that allows the system to very rapidly pick the optimal route by suppressing the decompositional route using CAR when possible, even before the second half of the compound is processed. In contrast, 6-7 year olds rely heavily on the decompositional route, where each constituent has to be evaluated as a single word.

Poster Session 1

Thee, uhh, role of discourse status in three-year-olds' understanding of disfluent utterances

Sarah J. Owens & Susan A. Graham (University of Calgary)
sjcollin@ucalgary.ca

Filled pauses are a common phenomenon in language. Despite the prevalence of filled pauses within speech, they have traditionally been characterized as an unnecessary divergence from speech that listeners must overcome (Clark, 1996). Recently, an alternative understanding has emerged. Specifically, because filled pauses are systematically produced before unfamiliar words (Arnold et al., 2007), words new to a discourse (Arnold & Tanenhaus, 2011), and in situations characterized by uncertainty (Brennan & Williams, 1995) or choice (Schachter et al., 1991), they are increasingly recognized as an informative element of the speech stream.

Research has demonstrated that adult listeners use disfluencies as a cue that a speaker is about to refer to something inaccessible or new to the discourse. In turn, this narrows the referential domain and facilitates reference resolution (e.g., Arnold et al., 2007). To date, only one study has examined the disfluency phenomenon in children. Specifically, Kidd et al. (2011) demonstrated that 32-month-olds' looks to unfamiliar discourse-new objects increased when an utterance was disfluent versus when the utterance was fluent. In this study, discourse status and object novelty were jointly manipulated. Research with adults, on the other hand, has demonstrated that adults use disfluencies as a referential cue based on either of the two factors independently. Thus, questions remain surrounding the precise nature and robustness of the disfluency effect in children. The current study addresses one such question by isolating the effect of discourse status. The research by Kidd et al. suggested emergence of the disfluency effect around 30-months, and because this task is complicated by the removal of one of the cues (object novelty), we tested 3 ½ year-old children.

Using an eye-tracking paradigm, 42-month-old children were presented with 16 familiar object pairs, each of which appeared on screen three times. The first two presentations established one object as *discourse-given*: children heard a recorded utterance labeling the object twice (e.g., "*I see the cat!*"). The other object (e.g., a horse) was not labeled, and was thus considered to be *discourse-new*. During the third presentation, children heard the critical instruction to look at either the discourse-given (e.g., the cat) or discourse-new (e.g., the horse) object. The instruction was fluent (i.e., contained no disfluencies: "*Look at the X*") or disfluent (i.e., "*Look at thee, uhh X*"). We analyzed children's looking times during two 400ms windows following onset of the noun; these windows were selected because they contain identical linguistic information regardless of fluency.

Results revealed a significant interaction between discourse status and fluency during the first 400 ms of the noun (*p* < .05). Specifically, in this initial stage of processing, children showed different eye gaze patterns depending on fluency of the utterance. During *disfluent* trials, children looked significantly more to the discourse-new object regardless of which referent was the target of the noun (either discourse-new or discourse-given, During *fluent* trials, however, children showed early evidence of processing the noun. That is, children demonstrated a higher proportion of looking at the discourse new object when the new object was the target, and a higher proportion of looking at the discourse given object when the given object was the target (p < .05). These results indicate that hearing "*thee uhh*" led children to anticipate reference to the discourse-new object, and that the disfluency disrupted children's ability to identify the target.   As predicted, this effect begins to disappear as the noun unfolds. These differences were present during the early stages of processing but, as expected, disappeared as children processed the semantic content of the noun (i.e., the second 400ms window following onset of the noun). Results will be discussed in the context of children's ability to use paralinguistic cues to facilitate referential communication.

Processing effects on grammar acquisition: Evidence from an artificial language study

Lucia Pozzan, Lila Gleitman & John Trueswell (University of Pennsylvania)
lpozzan@sas.upenn.edu

Introduction. A cross-linguistically reliable strategy to map sentences to interpretations is to posit as many distinct thematic roles as noun phrases (NP's) in a sentence[1]. This strategy is not error-free, and will lead learners to incorrect interpretations, if conflicting language-specific cues to argument structure (e.g., morphology) are overlooked. Here we explore how online parsing constraints, and in particular the known inherent difficulty in revising initial parses, may shape the ability to *acquire* and *use* morphological cues to argument structure. Evidence for this hypothesis exist in the literature: children learning Tagalog, a verb-initial language in which causative verb morphology is a reliable predictor of argument structure, show greater sensitivity to causative morphology than children learning Kannada, a verb-final, morphology-final language, perhaps because verb morphology can be used to guide parsing in verb-initial languages but can only be used to revise and confirm parsing in verb-final languages[2]. These two languages, however, have vastly different grammars and case marking systems. Moreover, children's comprehension, not production, of morphology was surveyed, leaving open the possibility that the observed differences reflect known difficulties in using morphology to revise initial interpretations in a morphology-final language and not parsing effects on grammatical acquisition *per se*.

Experiments. Here we report the results of an artificial language learning study demonstrating that both comprehension and production of morphology are delayed when morphological cues to argument structure appear at the end, rather than at the beginning, of sentences, in otherwise identical grammatical systems. Adults [N=36] learned one of four 'miniature' language variants over a three-day period. They were explicitly taught noun meanings but had to learn word order and morphology by observing actors carrying out actions in response to spoken instructions in the language (e.g., *Shake the pig with the tweezers*). In all language variants, morphology (indicated below as $_{MORPH}$) is a perfectly reliable predictor of argument structure (indicating the presence of 1 vs. 2 thematic roles). On the contrary, since all language variants allow (optional) argument omission, the number of NP's in a sentence (1 vs. 2) is not a reliable predictor of argument structure. In two of the languages, morphology appeared on the verb, and the verb appeared either sentence-initially (V_{MORP}-(NP)-NP) or sentence-finally (NP-(NP)-V_{MORP}). Two additional language variants, in which morphology appeared on NPs, were created to control for potential effects of word-order similarity between English and the verb-initial variant: a verb-final, morphology-initial (NP_{MORPH}-(NP)-V) variant, and a verb-initial, morphology-final (V-(NP)-NP_{MORPH}) one.

Results & Discussion. In both a comprehension task (act out) and a production task (describing actions) on Day 3, participants who learned a language where morphology appeared sentence-initially showed higher accuracy in their use of morphological cues than those learning a language where morphology appeared sentence-finally (p's<0.05), independent of whether morphology was on verbs or nouns (p's>0.5). (Results based on multi-level mixed effects logistic regressions, see Figure for average proportion correct.) These results provide strong evidence that on-line processing constraints impact language acquisition. Morphological cues that arrive early in the sentence, and as such *guide* interpretation, are easier to learn that late-arriving cues used to *revise* intitial interpretative commitments. Parallel performance patterns in production and comprehension indicates that parsing affects grammatical acquisition.
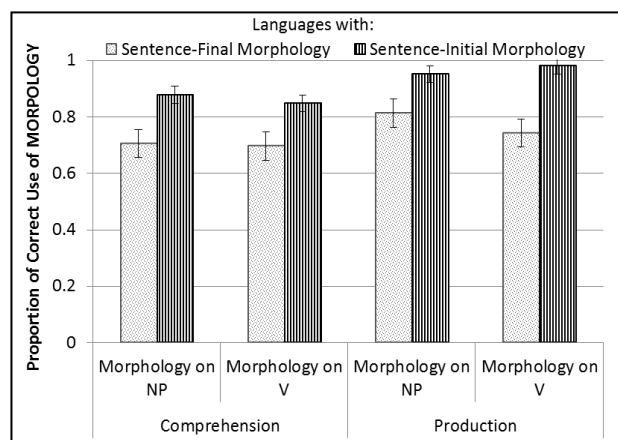
References. [1]Naigles, L., Gleitman, L.R., & Gleitman, H. (1993). *Language and cognition: A developmental perspective.* [2]Trueswell, J.C., Kaufman, D., Hafri, A. & Lidz, J. (2012). *Proceedings of BUCLD 36.*

Poster Session 1

Can anaphoric dependencies be primed across languages? Evidence from Italian-English bilinguals

Emily Fedele, Elsi Kaiser, Maria Zubizarreta (University of Southern California)
emilyfedele@gmail.com

A key question in bilingualism research is how the two languages are organized. Prior work has found evidence for shared syntactic and lexical representations across languages (e.g. Hartsuiker et al 2004). We used priming to test whether abstract representations activated during *pronoun resolution* are shared across languages, a question not previously tested (cf. Serratrice'07). Prior work on English found that interpretation of object pronouns is primed by anaphoric dependencies in preceding sentences (within-language effect, Kaiser 2009). Our **Exp#1** tested English speakers (n=16) to see if interpretation of ambiguous subject pronouns in English is influenced by anaphoric dependencies in preceding English primes. **Exp#2** tested Italian-English bilinguals (n=20; L1 Italian, high fluency in English) to see if interpretation of ambiguous subject pronouns in English is influenced by anaphoric dependencies in preceding *Italian* sentences. Italian has null and overt pronouns. Nulls tend to refer to preceding subjects and overts to objects (e.g. Carminati 2002).

Exp#1 primes used gender cues, and had subject-referring pronouns (1a), object-referring pronouns (1b), ambiguous pronouns (1c) or no pronouns (1d). **Exp#2** primes (in Italian) manipulated pronoun form (null/overt) and referents' gender (same/diff gender): Primes had (i) a form cue pointing to the preceding subject/object (2a-b), or (ii) a gender cue *and* form cue (2c-d), or no pronoun (2e). (Verb ending signals gender in (2c)). **Targets** were identical in Exp1/2, in English, with ambiguous subject pronouns (ex.3). Primes and targets were followed by questions on pronoun interpretation (*Who returned from the trip?).*

Exp#1: Primes in English
(1a) <u>Kate</u> picked flowers with Mark in a field when <u>she</u> slipped on a loose rock.　　[DiffGend/Subj]
(1b) Mark picked flowers with <u>Kate</u> in a field when <u>she</u> slipped on a loose rock　　[DiffGend//Obj]
(1c) Kate picked flowers with Mary in a field when <u>she</u> slipped on a loose rock.　　[SameGen/Amb]
(1d) Kate picked flowers with Mark in a field　　[baseline, no pronoun]

Exp#2: Primes in Italian (shown in English for convenience, actual sentences were in Italian)
(2a) <u>Rita</u> picked flowers with Maria in a field when <u>ø</u> slipped on a loose rock.　　[Same+null]
(2b) Rita picked flowers with <u>Maria</u> in a field when <u>she</u> slipped on a loose rock.　　[Same+overt]
(2c) <u>Maria</u> picked flowers with Paolo in a field when <u>ø</u> slipped on a loose rock.　　[Diff+null]
(2d) Paolo picked flowers with <u>Maria</u> in a field when <u>she</u> slipped on a loose rock.　[Diff+overt]
(2e) Maria picked flowers with Rita in the field.　　[baseline, no pronoun]

(3) **Exp#1/2: Targets in English**: Bill visited Fred when <u>he</u> returned from the trip. [ambiguous]

Results/Exp#1 (Eng-Eng): We extend Kaiser's findings to subject pronouns: All four conditions show a subject preference for the ambiguous pronouns in the targets (approx.61% subj-choices), but this preference is *significantly weakened after object-biasing primes* (ex.1b, 52%, p's<.05): Anaphoric dependencies from primes can influence ambiguous pronoun interpretation. Can this priming reach across languages? **Results/Exp#2 (Ital-Eng)**: Even when we focus on just those trials where the primes were interpreted as intended (i.e. people chose subject in (3a,c), object in (3b,d)), we find *no priming* in targets: Responses for the ambiguous pronouns in targets are split between subject and object in all conditions (no sig preference for subject or object, no sig differences between conditions), regardless of prime. Thus, we do not find clear evidence for priming.

Pronoun interpretation is primed by preceding anaphoric dependencies within a language (Exp#1), but not across languages (Exp#2). Given earlier findings that cross-language priming is absent when representations differ across languages (Loebell/Bock 2003, Bernolet et al 2007 on syntactic priming), this suggests the lack of cross-language pronoun-interpretation priming may stem from the different anaphoric paradigms of English and Italian. This has implications for the representations activated when processing anaphoric dependencies.

Children's and adults' processing of coordination ambiguity: An eye-tracking study

Justine VanDyke-Lyon[1], Lap-Ching Keung[2], & Fernanda Ferreira[1] ([1] University of South Carolina, [2] University of North Carolina – Chapel Hill)
vandykej@mailbox.sc.edu

Many sentences are temporarily ambiguous. For example, the conjunction *and* can connect two noun phrases as in (1), or two clauses as in (2). (1) *The girl knocked over the salt and the pepper after losing her balance;* (2) *The girl knocked over the salt and the pepper went flying across the table.* Previous research reveals longer reading times on disambiguating words (e.g., *went*) in sentences such as (2) compared to controls, suggesting that adults use frequency information to build expectations about likely syntactic structure; in English, the conjunction *and* conjoins NPs more often than clauses. Here we extend this line of research by asking whether this frequency effect is further influenced by expectations concerning the order in which NPs co-occur. To this end, we recorded participants' eye-movements as they read sentences containing coordinated noun phrases (binomials) that frequently co-occur in a particular order to binomials in the less frequent order (*The girl knocked over the pepper and the salt went flying across the table).* We manipulated ambiguity by including sentences with and without commas. We hypothesized that ambiguous, frequent binomials would be harder to process than their reversed counterparts because comprehenders are more committed to the highly expected, initial NP coordination interpretation. Additionally, we sought to investigate these same questions among adolescents. This is of particular interest since children's semantic and syntactic knowledge is known to be positively correlated with off-line reading comprehension. To date, few studies have investigated how this information is engaged in real-time by children as they are reading. Thus, in the current study we examined children's vocabulary knowledge as a predictor of sensitivity to syntactic (*Comma*) and collocation frequency effects (*Order*), as measured by reading times.

Twenty children (mean=11.5yrs) and 20 adults (mean=18.8yrs) read 48 experimental sentences and responded to a YES/NO question after each. The design was fully counter-balanced; filler trials ensured a balance of structures and YES/NO responses.

Offline comprehension: Overall, children made significantly more errors (mean accuracy=76%) compared to adults (91%). Adults were less accurate answering questions following binomials in the reversed order. Despite children demonstrating a similar pattern, differences among conditions were not significantly when random by participant variability was controlled.

Online processing: For adults and children, linear mixed effects models revealed a significant effect of *Comma* in the disambiguating region. Specifically, we observed longer reading times for ambiguous sentences compared to sentences with commas (garden-path (GP) effect). In addition, immediately following the disambiguating region (spill-over region), we observed significant effects of both *Comma* and *Order*. Adults and children demonstrated longer total reading times when there was no comma and when the binomials were in the reversed order. For adults, there was also a significant interaction such that total reading times were fastest when binomials were in the standard order and when there was a comma. These results suggest that two factors contributed to sentence processing difficulty: encountering a sequence of NPs in an unexpected order, and experiencing a garden-path.

Vocabulary knowledge as a predictor: Children's vocabulary scores were associated with GP effects in both the disambiguating and spill-over regions. As vocabulary scores increased, reading times in the ambiguous, no comma condition also increased. These results suggest that children with richer vocabularies have more robust syntactic expectations than their peers with weaker vocabularies. Vocabulary scores did not predict effects associated with the *Order* manipulation, suggesting that vocabulary measures are not sensitive enough to capture the subtle effects of co-occurrence frequency observed.

In summary, both children and adults were affected by two separate violations of their expectations: order of constituents and overall syntactic form. In addition, the vocabulary data from the children show the rich connection between lexical knowledge and syntactic parsing strategies.

Poster Session 1

The online processing of the Japanese anaphoric expressions *zibun-zisin* and *kare*

Atsushi Yuhaku (Ritsumeikan University), & Satoru Nakai (Doshisha University)
yuhakun@gmail.com

Introduction: Great attention has been paid to the online processing of anaphoric expressions in Indo-European languages [1], but little is known about the processing of anaphoric expressions in non-Indo-European languages. The purpose of this research is to examine how Japanese anaphoric expressions are processed online and to show that those processing is based on the processing hierarchy [2], which predicts that the processing cost of anaphoric expressions is linearized: the syntactic module < the semantic module < the discourse module.

Method: 23 Japanese-native university students (M=19.4 yrs old) joined the study. The experimental paradigm is the Cross-Modal Lexical Decision Method (CMLD). CMLD requires the participants to do the two tasks simultaneously. They listened to the stimuli and answered the questions (the primary task). At the same time, they had to judge whether the visually presented words were acceptable or not while listening to the stimuli (the secondary task). CMLD assumes that the more difficult the primary task is, the longer the response time (RT) for the secondary task is. Two types of stimuli were presented to the participants. The first type includes the Japanese reflexive *zibun-zisin* 'self-self' and the second type the Japanese pronoun *kare* 'he'. (1) [Refl] *Ekicho$_1$ wa Dezaina$_2$ ga mukasikara itumo zibun-zisin$_{*1/2}$ no yuujin o sinjiteita to itta.* 'The station master$_1$ said the designer$_2$ always trusted his$_{*1/2}$ friend.' (2) [Pron] *Ekicho$_1$ wa Dezaina$_2$ ga mukasikara itumo kare$_{1/2}$ no yuujin o sinjiteita to itta.* 'The station master$_1$ said the designer$_2$ always trusted his$_{1/2}$ friend.' In the above hierarchy, when an anaphoric expression and its antecedent are not the coarguments of the same predicate, but the bound variable relationship is formed, it is processed in the semantic module. On the other hand, when the bound variable relationship is not formed (covaluation), the anaphoric expression is processed in the discourse module. If the above processing hierarchy is correct, the reflexive *zibun-zisin* is to be processed in the semantic module and its processing cost is lower than that of the pronoun *kare*, which does not accept the bound variable relationship with the antecedent [3] and is to be processed in the discourse module.

Results: The accuracy rate for the primary task was 86.9% for the Refl and 80.4% for the Pron and there was no significant difference. In contrast, there was a significant difference (p=.025) in the RT for the secondary task between the Refl and the Pron. The participants responded much faster to the Refl (825.9ms) than to the Pron (926.1 ms) immediately *after* those anaphoric expressions, but no significant difference in the RT between the Refl (863.3ms) and the Pron (891.7 ms) at immediately *before* those expressions.

Discussion: These results showed that the participants processed the reflexive *zibun-zisin* much easier than the pronoun *kare*. First, let us consider the difference in the number of possible antecedents between the reflexive and the pronoun. As shown in the examples (1) and (2), the pronoun can refer to both the matrix subject and the embedded subject while the reflexive refers to only the embedded subject, which might cause the Pron to be more difficult. The primary task, however, served to confirm how the participants interpreted the anaphoric expressions and its data showed that they correctly chose the embedded subject as the antecedent for both the reflexive and the pronoun, so this possibility can be ruled out. Next, let us consider the above processing hierarchy. This hierarchy predicts that the reflexive in the non-coargument position is processed in the semantic module and that its processing requires less processing resources than the pronoun processed in the discourse module. This prediction is borne out.

Conclusion: This research shows that Japanese native speakers process the reflexive *zibun-zisin* much faster than the pronoun *kare*. This is in accordance with the prediction of the processing hierarchy of anaphoric expressions and indicates the application of the hierarchy in the online sentence processing of Japanese anaphoric expressions.

References: [1] Koornneef (2008) *Eye-catching Anaphora.* [2] Reuland (2001) Primitives of Binding. *L.I.*, 32, 439-492. [3] Hoji (1991). *Kare.* In Georgopoulos & Ishihara (Eds), 287-304.

Resolving temporary referential ambiguity using presupposed content

Jacopo Romoli (Macquarie University), Manizeh Khan (Harvard University), Yasutada Sudo (CNRS-EHEES-ENS), & Jesse Snedeker (Harvard University)
jacopo.romoli@mq.edu.au

Presupposed information is an important component of meaning arising at the interface between semantics and pragmatics. Presuppositions carry information that can potentially be used incrementally to guide interpretation during language processing; however, little is currently known about the processing of presuppositions (but see Schwarz, 2007; Schwarz & Tiemann, 2012). We present the results of two visual-world experiments investigating whether the presupposition of "also" is used to predict upcoming linguistic material during sentence comprehension. We compare predictions from "also" to predictions from "only".

Experiment 1: On each trial, participants (N=32) heard two context-setting sentences followed by a target instruction. In a 2x2 design, the target instruction either referred to a previously mentioned item or a new item (Old/New), and either included "also"/"only", as appropriate, or neither (Critical/Control). Each 4-item display contained two images depicting cohort-competitors (e.g. *candy, candles),* creating a temporary referential ambiguity during the target noun. A typical trial looked as in the following examples.

Figure 1. Log-odds of proportion looking to target versus cohort competitor.
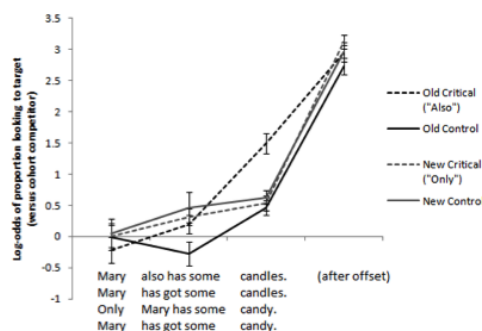
Context sentences: Michael and Sarah are friends, Michael has candles and shoes.

Target sentences (four conditions):

Sarah also has some candles (Old/Critical)

Sarah has got some candles (Old/Control)

Only Sarah has some candles (New/Critical)

Sarah has got some candles (New/Control)

A linear mixed effects regression on the log-odds of the proportion of looks to the target during the Noun yielded a significant main effect of Critical vs. Control (t=2.9,p<.01) and a marginally significant effect of Also vs. Only (t=1.9,p=.06). Critically, there was a significant interaction (t=3.6,p<.01), driven by greater anticipatory looking in the "also" critical condition but not in the "only" critical condition.

Experiment 2: we followed up on Experiment 1 to" a) test if "only" would be used predictively in a more felicitous context where there is another character with whom the speaker could be drawing a contrast and b) explore whether implicit mention would be sufficient to satisfy the presupposition of "also", by changing the context sentence (e.g. "Look at what Michael has"). Implicit vs. explicit mention was manipulated between-subjects (N=32 for each condition). Looks to the target during the Noun region showed a main effect of mentioning "also" (t=5.2,p<.01) but not of "only" (t<0.9,p>.3), and no interaction with the context manipulation in either condition (ps>.3). Further, analyses on earlier regions of the sentence indicated that anticipatory looks in the "also" condition began at the earliest possible moments, within 200ms of the offset of the word "also".

Discussion: Adults rapidly used the presupposition of "also" incrementally in online sentence comprehension. On the other hand, we saw no use of the inference that could be generated by hearing "only" (i.e. that the upcoming material will be unique). This is surprising because, in an offline sentence completion task that mirrored Experiment 1, participants were able to generate and use this inference. The present study confirms Schwarz's (2007) finding that presuppositions are processed online. Critically, by using a more temporally-sensitive design we are able to demonstrate that these effects emerge within 200ms of the offset of the presuppositional trigger.

Coherence expectations underlie parallelism effects for conjoined clauses

Laura Kertz and Corey Cusimano (Brown University)
laura_kertz@brown.edu

The source of **parallel facilitation** observed for conjoined clauses (Frazier et al. 1984) has been much debated, with competing explanations invoking either a general priming mechanism or a processing constraint unique to coordinate structures. In the current study we tested the hypothesis that parallel facilitation follows from constraints on information packaging that support **discourse coherence.** Specifically, we proposed that when a discourse highlights the resemblance between two events, as opposed, for example, to a causal relation, there is a strong preference to mention event participants in parallel order (cf. Kehler 2002). This ordering preference, we propose, underlies facilitation for clause-level parallelism. We tested this hypothesis in two ways.

In a **sentence completion** task (n=36) a prompt appeared in either active or passive voice followed by a connective biased toward either a resemblance ('while') or causal ('because') coherence relation (1-2 below). Overall, there was a strong tendency to supply completions in active voice (92%). However, for 'while' prompts, the rate of active completions depended on the form of the matrix clause (99% following an active; 79% following a passive). For 'because' prompts, there was no effect of voice on the rate of active completions (96% vs. 93%). This result indicates that preferences for parallel structure are conditioned on expectations regarding discourse coherence.

Next, we used a **visual world** paradigm (n=30) to test whether connectives interact with structure to bias listener expectations for upcoming referents. We predicted that, upon hearing 'while', listeners would launch anticipatory looks to a new discourse referent that is semantically similar to the subject of the preceding clause (e.g. 'detective' is more similar to 'cop' than 'biker'), but that looks following 'because' would be unaffected by the structure of the previous clause. We tested this with stimuli which described the same event in active or passive voice, followed by either 'while' or 'because' (3-4 below). We then tracked looks to a target, which was displayed along with three competitors.

Study 1: Sentence completion prompts

1. The nurse was examining the woman while/because the _____.
2. The woman was being examined by the nurse while/because the _____.

Study 2: Auditory stimuli

3. The cop was arresting the biker while/because the detective was subduing the bartender.
4. The biker was being arrested by the cop while/because the detective was subduing the bartender.
(**Visual target:** detective, **competitors:** cop, biker, bartender)

Growth curve analysis tracking looks to the target from the onset of the subject of the second clause revealed the predicted interaction between voice and connective ($p < .05$). Time bin analyses also confirmed that after hearing 'while', participants looked to a new referent semantically similar to the subject of the previous clause; after 'because' looks to previously mentioned participants persisted (presumably because listeners anticipated an explanation implicating these discourse-old referents).

These results demonstrate that sentence structure can guide expectations about order of mention for upcoming referents, specifically in cases where the listener is expecting a resemblance coherence relationship between the two clauses. The findings support an alternative view of parallelism effects in conjoined clauses as a special case of the influence of discourse-based expectation.

References. Frazier, L., Taft, L., Roeper, T., Clifton, J., & Ehrlich, K. (1984). Parallel structure: A source of facilitation in sentence comprehension. Memory and Cognition, 12(5), 421–430. Kehler, A. (2002). Coherence, Reference, and the Theory of Grammar. Stanford: CSLI.

Effects of event-structure and topic/focus-marking on pronoun reference in Korean
Kitaek Kim, Theres Grüter & Amy J. Schafer (University of Hawai'i at Mānoa)
kitaek@hawaii.edu

Reference resolution is known to be subject to a variety of interacting, and potentially form-specific, constraints (e.g., Arnold, 2001; Kaiser & Trueswell, 2008). Ueno & Kehler (2010, henceforth U&K) recently extended this line of research by asking (i) whether *null pronouns* in Japanese display the same interactions between grammatical and pragmatic factors as those observed for overt pronouns in English (Kehler et al., 2008), and (ii) whether topic-marking on the source argument of a transfer-of-possession event (see 2) would make that referent a more likely topic for the next sentence in the discourse. Their results from a passage completion task adapted from Rohde et al. (2006) suggested that (i) only overt *but not null* pronouns were sensitive to pragmatic factors in Japanese, and (ii) topic-marking did not influence reference resolution. Here we reconsider these claims in two experiments using closely related materials *in Korean*.

Exp1. Thirty native Korean speakers participated in a written passage completion task with a 2x3 design, varying *aspect* in the context sentence (perfective/imperfective) and *prompt type* in the continuation (null pronoun prompt/overt pronoun/free (i.e., unrestricted form)). Unlike in U&K, the null-pronoun prompt was not presented as "subject-omission", but as "Ø", explained as "invisible pronoun", to avoid creating a subject/source bias. Results, in line with previous work, show the subject of the continuation referred to the source of the context sentence more often when that sentence was imperfective, i.e., described an incomplete event, vs. perfective (Fig.1; main effect of *aspect*: $F_1(1,29)=14.5$, $p<.01$; $F_2(1,41)=12.0$, $p<.01$). Yet unlike in U&K, this effect did *not* interact with prompt type, suggesting that overt *and null* pronouns in Korean are sensitive to event-structure. We attribute this change to the less biasing prompt description.
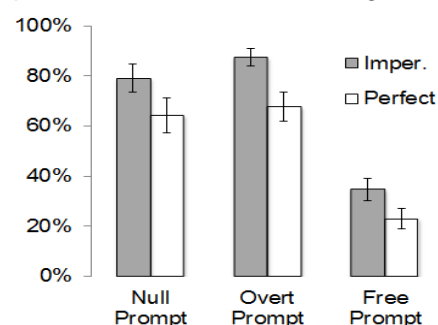
Fig.1. %Source-reference by prompt type

Exp2. The Korean marker *-nun* can co-occur with the source (2) or goal argument (3). Sohn (1999) describes –*nun* as a contrastive topic marker ("TC"). We predicted that in sentence-initial position (2), *-nun* would denote a topic whereas medial *-nun* would mark contrastive focus. Arguments marked with Topic-*nun* should be likely referents for subsequent pronouns. However, we predicted that ContrastiveFocus-*nun*, on the Goal argument, would promote continuations that maintained the Source argument as the topic and selected a contrastive Goal. Thus, Source-reference would be *higher* for (3) than (1).

(1) Chelswu-ka   Yengswu-eykey   chayk-ul   cwu-ess-e.
    Chelswu-Nom Yengswu-Dat   book-Acc give-Past   "C. gave a book to Y."
(2) Chelswu-**_nun_** Yengswu-eykey   chayk-ul   cwu-ess-e.
    Chelswu-**TC**   Yengswu-Dat   book-Acc give-Past   "**C.TC** gave a book to Y."
(3) Chelswu-ka   Yengswu-eykey-**_nun_** chayk-ul cwu-ess-e.
    Chelswu-Nom Yengswu-Dat-**TC**   book-Acc give-Past   "C. gave a book to **Y.TC**."

Thirty six native Korean speakers participated in a passage completion task with a 3x3 design, varying *TC-marking* in the context sentence (nominative-marking/Topic-*nun*/ContFocus-*nun*) and *prompt type* in the continuation (as in Exp1). Results show the subject of the continuation (for all three prompt types) referred to the source of the context sentence more often when that sentence had ContFocus-*nun* (65.7% Source-reference for (3)), but like U&K, no significant difference was found between nominative- and Topic-*nun*-marking (43.8% & 51.0% Source-reference for (1) and (2) respectively; main effect of TC-marking: $F_1(2,70)=18.5$ $p<.01$, $F_2(2,88)=15.658$, $p<.01$). Notably, contrastive continuations (e.g., *Chelswu did not give a book to MINHO*) were produced exclusively in ContFocus-*nun* conditions, constituting 29.1% of all produced sentences there, suggesting (contrastive) focus promotes alternatives in the reference set.

Clause structure matters: The role of left dislocations and clefts in pronoun resolution

Barbara Hemforth, Israel de la Fuente (Université Paris 7-Diderot), Saveria Colonna (Université Paris 8-Vincennes-Saint-Denis) & Sarah Schimke (Universität Osnabrück)
barbara.hemforth@linguist.univ-paris-diderot.fr

Pronoun resolution is essential for the interpretation of ongoing discourse and for establishing coherence. Therefore, it would be reasonable to assume that pronouns show a general preference for the most relevant or the most salient antecedent in the current model of the situation (e.g., Ariel, 1988; Chafe, 1974; Grosz et al., 1995). Consistent with this assumption, a number of studies have shown that potential antecedents increase in accessibility when they are established discourse topics, when they are left-dislocated and, presumably topicalized, or when they are foregrounded in a cleft construction (e.g., Cowles et al., 2007; Ellert, 2010; Foraker & McElree, 2007). Based on a series of visual-world studies and off-line questionnaires, we will show that, while this assumption seems to be generally in line with pronoun resolution across sentence boundaries (1a), within-sentence pronoun resolution, where the potential antecedents are in the matrix clause and the pronoun in a subordinate clause (1b), seems to contradict this general relevance or salience principle.

(1a) As for the postman he… / It was the postman who… / The postman… / …met the streetsweeper. He was on his way home.
(1b) As for the postman he… / It was the postman who… / The postman… / …met the streetsweeper when he was on his way home.

In particular, our results consistently show that, while the ambiguous pronouns in (1a) and in (1b) show an increased preference for left-dislocated antecedents, clefting enhances accessibility only between sentences. Within the sentence, however, clefting seems to have the opposite effect, that is, it decreases the accessibility of the clefted antecedent for ambiguous pronouns in the subordinate clause. We will show that this pattern holds for overt subject pronouns in French and in German, as well as for null subject and object clitic pronouns in Spanish. While basic preferences in sentences with SVO matrix clauses differ considerably and predictably across constructions and languages, left-dislocation and clefting generally show a highly similar effect increasing (for left-dislocations) or decreasing (for clefting) accessibility. Additionally, the results from a questionnaire in French suggest that this pattern holds even in cases where the cleft structure is used in the context of a narrow focus question ("Who met the street-sweeper?) and in the context of an all focus question ("What happened?"), where they are more likely interpreted as presentational clefts.

We will argue that the observed dispreference for the clefted antecedent in within-sentence pronoun resolution stems from properties of both the semantics and the information structure of cleft sentences. Cleft sentences consist of a foregrounded (potentially new, focused) part, the cleft, and a background (potentially given, presupposed) part, the coda. We assume that the different possible interpretations of the pronoun are due to different attachment preferences of the subordinate clause as a whole, which can be attached high (modifying the whole sentence) or low (as part of the presupposed information). A preference to interpret the subordinate clause as presupposed will explain the cross-linguistic and cross-construction results. We will argue more generally that coherence is primarily established on corresponding informational levels, i.e. within the backgrounded/presupposed or the foregrounded/asserted part of an utterance.

Online sensitivity to structural constraints on bound variable anaphora
Ian Cunnings (University of Edinburgh), Clare Patterson & Claudia Felser (University of Potsdam)
ian.cunnings@ed.ac.uk

A growing body of research has examined the parser's sensitivity to structural constraints on reference resolution. Much of this work has focused on reflexives, showing that structural constraints on binding (notably, c-command) affect parsing immediately [5, 6]. Comparatively few studies have looked at the processing of bound variable anaphora. Linguistic theory posits that for a quantifier phrase (QP) to bind a pronoun, the QP must c-command the pronoun [4]. Exceptions are however attested [1], and an earlier study [2] found no evidence of c-command failure increasing processing difficulty, although these findings might have resulted from the fact that the materials tested contained only one (gender matching) antecedent in the discourse. We report the results from a new eye-movement experiment, which investigates whether the c-command constraint on bound variable anaphora affects online sentence comprehension.

32 native English speakers read 24 critical items in a 2x2 design as in (1/2), plus 60 fillers. Gender congruence was manipulated between the QP and pronoun as a diagnostic for dependency formation. In the (a) versions of each sentence, the QP matches in gender with the pronoun, while in the (b) versions there is a gender mismatch. In each case, *the surgeon* is a grammatically available antecedent that matches in stereotypical gender with the pronoun.

(1a) The surgeon saw that every old man on the ward silently wished that he could go a little bit faster.
(1b) The surgeon saw that every old woman on the ward silently wished that he could go a little bit faster.
(2a) The surgeon who every old man on the ward saw silently wished that he could go a little bit faster.
(2b) The surgeon who every old woman on the ward saw silently wished that he could go a little bit faster.

The c-command requirement predicts that the pronoun *he* should be able to bind to the QP in (1), when the QP c-commands the pronoun, but not in (2), when it does not.

At the critical region (*that he*), we found reliably longer 1st pass times for sentences such as (2a,b) in comparison to (1a,b), indicating that sentences containing relative clauses were generally read slower than those without. In 2nd pass times at the pronoun region, we observed a reliable 'c-command' by 'gender' interaction. Planned comparisons revealed longer reading times (p = .017) in gender mismatch condition (1b) than in gender match condition (1a), but no reliable differences between (2a) and (2b) (p = .866). Reliable interactions were also observed at the spillover region (*could go*), where again second-pass times indicated longer reading times in (1b) than (1a) (p < .001), but no differences between (2a/b) (p = .842). A marginally significant interaction was also observed for regression paths at the sentence final region (*bit faster*), which again showed the same trend between conditions (1a,b) (p = .073) but not (2a,b) (p = .472).

Our finding of gender mismatch effects in (1a,b) but not (2a,b) indicates that participants only attempted to link the pronoun to the QP when it c-commanded the pronoun. These results contrast to a previous study [3] where non-quantified proper name antecedents inside relative clauses *were* considered as coreference antecedents for a subject pronoun. Together, these results suggest an antecedent search mechanism that is sensitive to structural constraints on both syntactically (e.g. reflexives) and semantically bound anaphora.

References

[1] Barker (2012). Linguistic Inquiry 43, 614-633; [2] Carminati et al. (2002) Journal of Semantics 19, 1-34; [3] Cunnings et al. (submitted). Variable binding & coreference in sentence comprehension; [4] Reinhart (1983). Anaphora & semantic interpretation; [5] Sturt (2003). Journal of Memory & Language 48, 542-562; [6] Xiang et al. (2009). Brain & Lang. 108, 40-55.

Syntactic prominence in the processing of reference: Does subordination matter?

Wei Cheng, Jenn Olejarczyk, Amit Almor (University of South Carolina)
cheng22@email.sc.edu

Repeated reference is one of the primary means to make text or conversation coherent. One important factor affecting the resolution of repeated reference is the prominence of the referent in the discourse, which affects the likely form used for the reference. For example, the repeated name penalty (the slower reading of repeated name anaphors than pronouns) occurs when the anaphor is coreferential with a prominent entity in local discourse, but disappears when the referent is in a non-prominent position. Referent prominence is determined by many factors, one of which is the syntactic role of the referent. Discourse Prominence Theory (DPT; Gordon & Hendrick, 1998) specifies that a referent's syntactic prominence is dependent on its height in the phrase structure, and is inversely related to its depth of embeddedness. Previous tests of this claim have looked at conjoined noun phrase subjects and possessive noun phrase subjects. Since embeddedness and prominence are confounded with constituent headedness in these constructions, it is not clear whether the results of previous studies support DPT's specific claims about the cause of prominence, or whether they simply reflect the unprivileged status of non-headedness of the referent. The present study aimed to disentangle these issues by comparing intersentential coreference with intrasentential coreference processing in constructions where the antecedent is embedded within subordinate clause(s) yet still occupies the subject head noun position, and thus should not be prominent according to DPT's analysis of prominence (see sample items).

Two self-paced reading experiments were conducted that differed in the degree of embeddedness of the referent (see sample items). Twenty-four and twenty-six native English speakers participated in Experiments 1 and 2 respectively. Both experiments had a 2×2 design with Structure (intra- vs. intersentential) and Anaphor (pronoun vs. repeated name) as independent variables. Each experiment had 32 test items constructed in four versions, which were counter-balanced so that each participant saw one version of each item. Participants read the stimulus clause-by-clause/sentence-by-sentence.

In Exp. 1, there was no main effect of Structure [F_1, $F_2 < 1$]. There was a main effect of Anaphor such that reading times were longer in the repeated name condition than the pronoun condition [$F_1(1, 21) = 7.38$, $p < .05$, $F_2(1, 31) = 6.30$, $p < .05$]. There was no significant interaction between the factors [$F_1 < 1$, $F_2(1, 31) = 3.00$, $p = .093$]. In Exp. 2, there was a main effect of Structure [$F_1 (1, 24) = 4.85$, $p < .05$, $F_2 (1, 31) = 6.51$, $p < .05$], with the second clause being read faster in the intrasentential condition, and a main effect of Anaphor, such that clauses/sentences containing pronouns were read faster than those containing repeated names [$F_1(1, 24) = 4.06$, $p = .055$, $F_2(1, 31) = 5.23$, $p < .05$]. There was no significant interaction between the factors [F_1, $F_2 < 1$]. We focus here on the main effect of Anaphor and lack of interaction found in both experiments.

Taken together, these results show that the repeated name penalty occurs when the antecedent is embedded within subordinate clause(s), a non-prominent position according to Discourse Prominence Theory. This indicates that syntactic prominence may not be determined only by the height of an antecedent in the syntactic tree and that its syntactic function (subject) and status (head noun) or topichood may play a more important role. The results also suggest that the same processes underlying coreference operate both within and between sentences.

Exp. 1 (intrasentential): When John went to the store, he/John saw a squirrel.
Exp. 1 (intersentential): John went to the store. He/John saw a squirrel.
Exp. 2 (intrasentential): Although it was said that Alex would come, he/Alex did not show up.
Exp. 2 (intersentential): It was said that Alex would come. He/Alex did not show up.

Reference

Gordon, P. C., & Hendrick, R. (1998). The representation and processing of coreference in discourse. *Cognitive Science, 22,* 389-424.

Figuring out Kafka: Structural biases induce early sense commitment for metonyms
Joel Fishbein & Jesse A. Harris (Pomona College)
jesse.harris@pomona.edu

According to the Underspecification Model, the language processor initially activates an underspecified representation of a metonym, consistent with all of its related senses (Frazier & Rayner, 1990; Frisson & Pickering, 1999, *et seq*). A specific sense may then be selected in a secondary *homing-in stage* if necessitated by context. Crucially, the model does not fully specify which types of information may activate the homing-in stage and at which point during sentence interpretation such information is relevant. Little research to date has investigated whether structural defaults – in addition to lexical and contextual information – can entice the parser to commit to a specific lexical sense (though see Fodor & Inoue, 2000). We investigated how sense selection was affected by an independently plausible principle, the *Subject-as-Agent Principle (SAP): provisionally assign an animate sentential subject an agent theta-role, all else being equal* (see also Grodzinksy, 1986, and Ferreira, 2003). With evidence from three studies on producer-for-product metonymies (e.g., when *Kafka* refers to the *works of Kafka*, rather than the literal, person sense), we propose that SAP encourages the processor to select a specific sense before required by context or lexical-level constraints.

Experiment 1: Effect of syntactic voice in reading. Experiment 1 was composed of two sub-experiments, a self-paced reading (N=32) and an eye movements (N=36) study, manipulating (i) Sentential-Voice (Active vs. Passive) and (ii) Verb-Type (Literal-selecting vs. Figurative-selecting; matched for frequency and length), as in (1). After both studies, we tested each subject's familiarity with metonymic names, discarding unfamiliar metonyms from the analysis. According to SAP, the processor provisionally assigns an agent theta-role to the metonym in Passive (*Kafka* was the subject; 1b), but not Active (*Kafka* was the object; 1a) voice. As an agent role is only consistent with the literal sense, the processor should engage in early sense selection for metonyms only in Passives. There were no processing differences for Actives, replicating previous results from Frisson and Pickering, and others. However, in the post-verbal region in Passives, items with Figurative-selecting verbs (*printed*) elicited slower reading times (M=752ms, SE=31) than Literal-selecting (*contacted*) counterparts (M=660ms, SE=37) in self-paced reading, t=2.09, p<0.05. In the eye movement follow-up, a processing cost for Figurative-selecting verbs over Literal-selecting ones was observed *only* for Passive items, manifesting in longer go-past times, increased regressions out, and longer second pass times on the post-verbal region, as well as increased second pass times and regressions in for the verb region.

Experiment 2: Fill-in-the-blanks. Items consisted of the Passive and Active sentence frames from Experiment 1, with the verbs replaced by blanks. The experiment was conducted over the Internet using Amazon Mechanical Turk. Participants (N=32) were instructed to provide the first sensical verb that came to mind. Even when biasing against our hypothesis by counting ambiguous verbs as metonymic, we found that fewer Figurative-selecting verbs were supplied for Passive (M=58.56%, SE=2.59) than for Active (M=71.82%, SE=2.37) sentences, z=3.30, p<0.001, as predicted, and various annotation schemes supported the results.

Experiment 3: Independent effect of context. To investigate the effect of prior context, this self-paced reading task (N=48) presented passives with metonymic subjects, manipulating (i) Verb-type as in Experiment 1, but varying the preceding context (ii) Contextutal-Bias (Figurative-Biasing vs. Neutral). There was again a processing cost for Figurative-selecting over Literal-selecting verbs (d=115ms), t=2.92, p<0.05, as well as facilitation for Figurative-Biasing over Neutral contexts (d=83), t=3.01, p<0.01, but no interaction, indicating that the processor was not tempted to override SAP, nor its earlier commitment to the literal sense, until it received information which was grammatically inconsistent with the initial sense selection.

Conclusion. The results above support the claim that a grammatical heuristic like SAP can motivate early commitment to a specific sense of a metonym and that application of this default is difficult to overturn by more general context; thus grammatical biases, as well as lexical-level constraints, may initiate an early commitment to a specific sense when processing metonyms.

(1) a. $|_1$ As planned, $|_2$ Kafka $|_3$ was printed/contacted $|_4$ by the publisher $|_5$ shortly after $|_6$ the revisions were in.
   b. $|_1$ As planned, $|_2$ the publisher $|_3$ printed/contacted $|_4$ Kafka $|_5$ shortly after $|_6$ the revisions were in.

Walking the walk and talking the talk, and perceptually simulating both while reading

Mallory C. Stites & Kiel Christianson (University of Illinois, Urbana-Champaign)
stites2@illinois.edu

Previous research shows that the semantic content of dialogue descriptions affects reading times on embedded quotes. Yao and Scheepers (2011) showed that readers were faster to read direct quotes when the preceding context implied a fast-talking character, an effect attributed to perceptual simulation of talker speed. Subsequently, Stites, Luke, and Christianson (2012) manipulated a character's action speed independently from their speaking rate (e.g., *John ran/walked into the room and said quickly/slowly, "I finally found my car keys"*) to determine if these two cues have separable effects on direct speech simulation. Readers spent less time reading direct (but not indirect) quotes explicitly described as being said quickly compared to slowly, whereas action speed had no effect. Furthermore, semantically "fast" adverbs were read faster than semantically "slow" adverbs, controlling for length and frequency.

However, Stites et al. (2012) leaves the open question of whether adverb speed affects quote reading times because it directly describes speech rate, or because it was always closer to the quote than the action verb was. In the current stimuli, the adverb describing speech rate is farther away from the direct quote than the verb (e.g., *John said quickly/slowly as he ran/walked into the room, "I finally found my car keys"*). The current results replicate Stites et al.'s findings that readers use adverb content to modulate direct quote reading times: go-past and total times are significantly shorter on quotes said quickly compared to slowly, with no effect of action speed. Thus, readers use the adverb describing the quote to guide simulation of talker speed, regardless of its physical distance from the direct quote. Additionally, fast adverbs were again read significantly faster than slow adverbs, with a marginal effect of adverb speed on the following phrase (*as he*), marking an important replication of the adverb effect in Stites et al.

Readers also appeared to perceptually simulate the speed at which a character performed a physical action, an effect that was not present in previous work. Total reading times on the action verb were shorter for fast than slow verbs (controlling for length and frequency), while selective go-past durations on the subsequent prepositional phrase describing the action performed (*into the room*) were shorter following fast relative to slow actions, with no effect of adverb speed on either region. To our knowledge, we are the first to show that perceptual simulation of a non-vocal action can modulate eye movements. This effect was not present in Stites et al. (2012) when the action occurred before the speech act was introduced, suggesting that the verb "said" cues the upcoming speech, causing readers to prepare to generate voice-related representations. This preparation for perceptual simulation could make readers more likely to create any perceptually-driven representations, including character movement rate. The dissociation between the effects of character speaking and movement rate suggests that readers generate perceptual representations that are sensitive enough to affect reading times on only the phrases to which they directly correspond, rather than generally affecting eye-movement speed over the remainder of the sentence.

Total Time for Regions of Interest
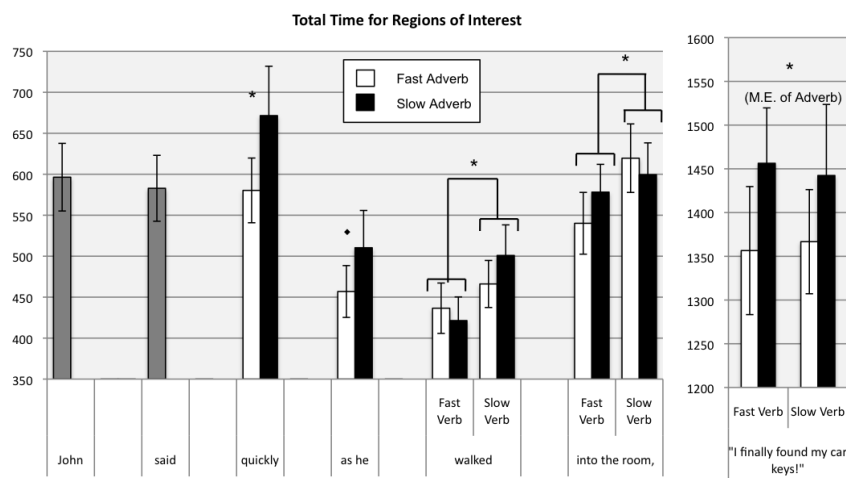
Can the bucket be kicked by him? – The processing of passivized idiomatic and literal sentences
Laura Dörre & Eva Smolka (University of Konstanz)
Laura.Doerre@uni-konstanz.de

Idioms like *"kick the bucket"* hold a special status in sentence processing. Their meaning cannot be constructed from the meaning of the individual parts. In addition, idioms are both semantically fixed, since single words cannot be exchanged, as in *"he punched the bucket"* and syntactically fixed, since they cannot undergo all syntactic transformations, as in *"the bucket was kicked by him"*. (e.g., Gibbs & Gonzales, 1985). In a previous study, we have shown that idioms are not as semantically fixed as previously assumed. Participants recognized the idiomatic meaning, even if single words were exchanged with semantic associations, as in *"she always reached for the planets"* or *"she always grasped at the stars"* (Smolka & Dörre, 2012).

The present study examined the syntactic fixedness of ambiguous idiomatic sentences, that is, sentences that hold both a figurative and literal meaning: Are idioms interpreted figuratively, even if they are passivized? In a sentence-completion test, we compared the recognition of active and passivized idiomatic and literal sentences The sentences were passivized in a way that the canonical passive word order was preserved.To this end, 52 idiomatic sentences were paired with 52 literal control sentences holding the same verb in the last sentence position. Half of the sentences were (a) transitive, the other half (b) ditransitive. In addition, 154 literal filler sentences were used.

(a)   *Der Redner hat den Rahmen gesprengt / Der Rahmen wurde vom Redner gesprengt.*
        Literal: The speaker has blown up the frame / The frame was blown up by the speaker.
        Figurative: The speaker went beyond the scope of his time.

(b)   *Sie hat ihm das Genick gebrochen / Ihm wurde von ihr das Genick gebrochen.*
        Literal: She has snapped his neck / His neck was snapped by her.
        Figurative: She brought him to ruin.

31 participants heard the sentences via headphones in either active or passive voice without the verb in the last sentence position. Three possible verbs were presented simultaneously on a screen, one completed the figurative meaning, one a semantically related meaning, and the third an unrelated meaning. Participants had to decide as fast as possible via a push button-box, which of the three presented verbs best completed the sentence. An error arose if the participants did not choose the verb that completed the figurative meaning. RT and error data showed easier processing of idiomatic than literal sentences, and of active than passive sentences. The lack of an interaction suggested that idiomatic sentences are recognized even if they are presented in passive voice. Interestingly, ditransitive passive sentences showed fewer errors than transitive passive sentences. This indicates that the passivization of idioms depends on syntactic rather than semantic information and that idioms are not as syntactically fixed as previously assumed. We integrate these findings in a model of idiom processing, with a focus on canonicity and adjacency of sentence constituents.

References

Gibbs, R.W. Jr. & Gonzales, G.P. (1985). Syntactic frozenness in processing and remembering idioms. *Cognition, 20*, 243-259.
Smolka, E., & Dörre, L. (2012, June). *Can You Reach for the Planets? – The Processing of Idioms in Aphasics.* Poster presented at the International Conference – NeuroPsychoLinguistic Perspectives on Aphasia, Toulouse, France.

Listen to the hand: Gestures shape the comprehension of ambiguous pronouns
Stephani Foraker & Megan Delo (SUNY Buffalo State)
forakesm@buffalostate.edu

The hand gestures people make while speaking are tightly coordinated with what they are saying (McNeill, 1992), and can communicate the meaning of nouns and verbs (Bernardis, Salillas & Caramelli, 2008) or the position and size of objects (Holler, Shovelton, & Beattie, 2009). We tested whether comprehenders use gestures to interpret pronouns. Speakers sometimes use the location or hand shape of gestures to help mark co-reference between a pronoun and its referent (Foraker, 2010; So, Kita, & Goldin-Meadow, 2009), but it is not clear when and to what degree comprehenders make use of gestures carrying co-reference information. One study by Goodrich Smith and Hudson-Kam (2012) indicated that gestures that are incongruent with order of mention in a spoken discourse can shift comprehenders' interpretation of a pronoun. In the present study, we tested for facilitatory rather than inhibitory effects of gesture on ambiguous pronoun interpretation. We predicted that a gesture consistently indicating one character should bias interpretation to that character, no matter its order of mention.

Comprehenders viewed 24 short videotaped discourses. As in the example below, the first part included a gesture accompanying each of two same-gender characters. In the critical last sentence during the pronoun, the gesture was (1) Person 1 repeated, (2) Person 2 repeated, (3) Ambiguous 2-handed symmetrical, or (4) No gesture. After viewing each video, the 32 native English-speaking participants answered a question probing pronoun interpretation.

Video: "Craig and Matt went on vacation. Craig$_{[Person\ 1\ gesture]}$ took a trip to Hawaii while Matt$_{[Person\ 2\ gesture]}$ took a trip to Florida. He$_{[1/2/3/4]}$ thought the weather was great while on vacation."
Question: Who thought the weather was great while on vacation?
Answer choices 1-7 scale: (1) definitely Craig … (4) either Craig or Matt … (7) definitely Matt

The Person 1 and 2 gestures were performed by different hands, counterbalanced across items. The gestures were either predominantly representational gestures with hand shape information or points with location information (consistent within an item). Properties of the speech across the four conditions were controlled so the speech alone did not bias interpretation, confirmed by a norming experiment of the audio extracted from each video condition, and items were counterbalanced across four lists.

Consistent with our predictions, comprehenders were sensitive to the speaker's gestures, using them to guide pronoun interpretation. In comparison to the Ambiguous gesture baseline condition, Person 1 gestures biased interpretation toward the first-mentioned character, and Person 2 gestures toward the second-mentioned character. The same pattern was found for comparisons with the No gesture baseline condition. Additionally, the baseline conditions differed: the No gesture condition showed a greater recency bias, while the Ambiguous gesture attracted ratings significantly closer to the ambiguous "either one" mid-point of the scale. There were no differences between representational and pointing gestures, $p = .15$.
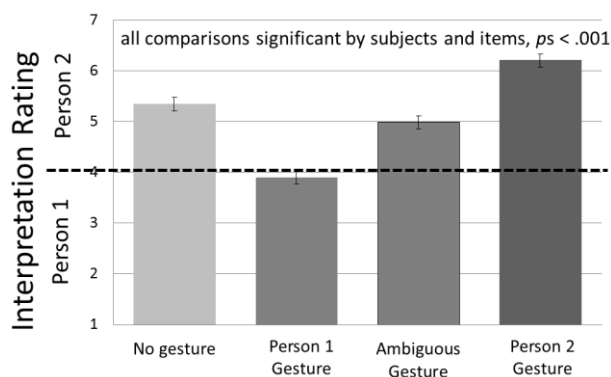
This experiment clearly shows that when gestures are present, comprehenders do use them to facilitate interpretation of a pronoun, utilizing the content of the gestures in relation to discourse structure. It remains to be seen whether such gestures are used in an automatic, obligatory fashion as part of on-line co-reference resolution, as Kelly, Ozyurek, and Maris (2010) have claimed for verb interpretation.

The Action-Sentence Compatibility Effect in American Sign Language

Kristen Secora (San Diego State University and University of California, San Diego) & Karen Emmorey (San Diego State University)
krsecora@gmail.com

Evidence from the embodied cognition literature suggests that humans engage the sensorimotor system during language comprehension and that one mechanism by which this occurs is mental simulation (Barsalou, 2008; Gallese & Lakoff, 2005). Effects of this simulation on motor execution have been demonstrated in stimulus-response compatibility effects such as the Action-sentence Compatibility Effect (ACE) (Glenberg & Kaschak, 2002). Response times (e.g., for sentence plausibility judgments) are facilitated when the motor response (e.g., pressing a 'yes' button that requires a movement toward the body) is congruent with the movement direction implied by the semantics of a written or spoken sentence (e.g., "You open the drawer"). Such facilitation provides evidence for the involvement of sensorimotor systems in language comprehension. In sign languages, however, there is a potential conflict between sensorimotor systems and linguistic semantics because movement away from the signer's body is perceived as motion toward the addressee's body (who is facing the signer). For example, semantics of the verb PUSH involve movement away from the body, but the addressee perceives the movement of the verb PUSH as toward, rather than away from their own body. Kaschak et al. (2005) found that perceiving nonlinguistic visual motion toward or away from the perceiver impacts the comprehension of motion sentences expressing motion toward or away from the body. We examined whether perceptual processing of linguistic sign movement modulates the ACE or whether the ACE is driven purely by the semantics of the verb. If the latter, then the direction of visual movement seen in the sign should have little effect on the motor response. If the former, then conflicting perceptual and semantic motion should modulate the ACE. Deaf ASL signers (N = 31) performed a plausibility judgment task for signed sentences that expressed motion away (e.g., "you threw a ball") or toward (e.g., "you grabbed a cup"). Responses were made by pressing a button requiring movement away from or toward the body. We found a significant congruency effect only when responses were categorized in relation to semantic motion rather than perceptual motion of the observed signed sentence [$F(1, 29) = 4.81$, $p<.05$]. This result indicates that a) the motor system is involved in comprehension of a visual-manual language and b) motor simulations for sign language are modulated by verb semantics rather than by the perceived visual motion of the hands.

References

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617-45.

Gallese, V. & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, *22*(3/4), 455-479.

Glenberg, A.M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin*, *9(3)*, 558-565.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., & Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition*, *94*, B79-B89.

Case-marking affects word order: Evidence from the gesture paradigm

Eunice Lim, Evelina Fedorenko, Edward Gibson (Massachusetts Institute of Technology)
eunjeeml@mit.edu

Although languages vary in the word order they use to describe transitive events (e.g., Dryer, 2002), Goldin-Meadow et al. (2008) have reported a universal word-order bias in a paradigm where participants are asked to convey event meanings using gestures. In particular, participants tend to gesture the verb (i.e., the action) after the subject and object (i.e., the event participants). For example, when asked to gesture the meaning of an animation depicting a boy kicking a ball, participants typically gesture "boy ball kick", regardless of whether they speak a verb-final language like Turkish, Japanese or Korean, or a verb-medial language like English, Spanish, Italian or Chinese (Goldin-Meadow et al., 2008; Langus & Nespor, 2010; Gibson et al., in press). Subsequent studies have, however, established that the default verb-final order can change to verb-medial for gesturing the meanings of reversible events such as a boy kicking a girl (Hall et al., 2010; Meir et al., 2010; Gibson et al., in press). Gibson et al. (in press) argued that the SVO order is more robust to noise because the positions of the noun phrases with respect to the verb can provide cues about which noun phrase is the subject and which the object (cf. the SOV order where both nouns are on the same side of the verb). Thus the loss of one of the noun phrases due to noise (e.g., Shannon, 1949; Levy et al., 2009) does not affect the interpretation of the remaining noun phrase in the SVO order: "girl kicks" allows the listener to recover the meaning of the girl kicking someone/something, and "kicks boy" allows the listener to recover the meaning of the boy being kicked.

However, if the SVO word order is more robust to noise, why aren't all languages verb-medial? Although SVO is the dominant order in a large fraction of the world's languages (41.2%), a similarly large fraction (47.1%) have SOV as the dominant order (e.g., Dryer, 2002). Gibson et al. (in press) have proposed that languages may retain the SOV order if they invent a system for marking the event participants for their thematic roles, such as case-marking or agreement. Indeed, in the gesturing paradigm some participants spontaneously use spatial cues to mark thematic roles (Hall et al., 2010; Gibson et al., in press). We report two studies that systematically investigate the use of such spatial "case-marking" and its effects on word order.

In Experiment 1 native English-speaking participants were shown animations depicting 16 transitive (8 non-reversible and 8 reversible) and 8 intransitive (filler) events. They first described each animation verbally, and then gestured each animation's meaning. Consistent with previous findings, participants i) uniformly used the SVO order in the verbal descriptions, ii) predominantly used the SOV order for gesturing non-reversible events (68% of trials), and iii) predominantly used the SVO order for gesturing reversible events (71% of trials). Critically, spontaneous spatial case-marking had a strong effect on the gesture ordering for the reversible events: 64% of the spatially-marked trials used SOV order (cf. 14% of the non-spatially-marked trials). In other words, the use of spatial case-marking made participants more likely to retain the default SOV order.

Experiment 2 used the same procedure as in Experiment 1, but included a third component where participants were explicitly instructed to use spatial case-marking: they were told to use their left hand to gesture the agent, and their right hand to gesture the patient. The results from the verbal and the initial gesturing (with no specific instructions) components replicated the results of Experiment 1 (65.5% SOV for non-reversible events, 69.2% SVO for reversible events). In the critical component (gesturing with case-marking instructions), participants retained the SOV order for both non-reversible (87.6%) and reversible (85.5%) events.

In conclusion, when case-marking information is available (generated either spontaneously or following specific instructions), participants retain the default SOV word order. This pattern of results provides strong support for the idea that at least two strategies exist in developing efficient communication systems. One strategy is to shift from the default SOV order to the SVO order to maximize meaning recoverability. And another strategy is to develop a system(e.g., case-marking) for indicating the thematic roles of the event participants.

Grammatical role primes spatial attention

Timothy W. Boiteau & Amit Almor (University of South Carolina)
twboiteau@gmail.com

Previous research has shown that numerical processing is tied with spatial attention, such that for native speakers of languages with left-to-right writing systems, numbers 1-4 prime responses towards the left side of space, and numbers 6-9 prime responses towards the right side of space [1]. We hypothesized that the canonical positions of nouns within subject-verb-object (SVO) sentences in English will similarly impact comprehenders' spatial attention. Thirty-seven participants read single transitive-sentence items, presented one word at a time on the center of the computer monitor. All items contained both a male and female character, counterbalanced across items to appear as Subject or Object (see Table). Following each sentence, one of the character names appeared on either the left or right side of the screen. Participants were instructed to judge the character's gender, counterbalancing button-gender assignment across participants. The dependent measure was response time (RT). They were told that following the experiment there would be a comprehension test about the items they read, encouraging them to read each item carefully. A 2 x 2 x 2 repeated measures ANOVA with factors Role (Subject vs. Object), Orientation (Left vs. Right), and Response Compatibility (Match vs. Mismatch) found significant effects for Role by Orientation, $F(1, 36) = 9.92$, $p = .003$, and Orientation by Response Compatibility, $F(1, 36) = 7.01$, $p = .012$. The Role by Orientation interaction was characterized by faster responses to Subjects appearing on the Left than on the Right, $t(36) = 2.46$, $p = .02$, and faster responses to Objects appearing on the right than to Subjects on the right, $t(36) = 2.39$, $p = .02$. The Orientation by Response Compatibility interaction was characterized by faster responses to the Left side of space than to the Right when response orientation and target orientation matched, $t(36) = 2.25$, $p = .03$ (see Figures). Our results show that grammatical role impacts spatial attention, probably as a result of repeatedly encountering written sentences with the canonical SVO format. Given the involvement of spatial mechanisms in referential processing [2] and the tight link between anaphora and gesturing [3], our findings may have implications for how referents are initially stored in the discourse model.

Timing	Event	Subject Left	Subject Right	Object Left	Object Right
300ms	Prime	Mary	Mary	Mary	Mary
300ms	Prime	kicked	kicked	kicked	kicked
300ms	Prime	John	John	John	John
300ms	Prime	in	in	in	in
300ms	Prime	the	the	the	the
300ms	Prime	face.	face.	face.	face.
500ms	Fixation	+	+	+	+
4000ms	Target	Mary	Mary	John	John

Figures. RT in milliseconds on the *y*-axis and target orientation on the *x*-axis. Bars represent *SE*.
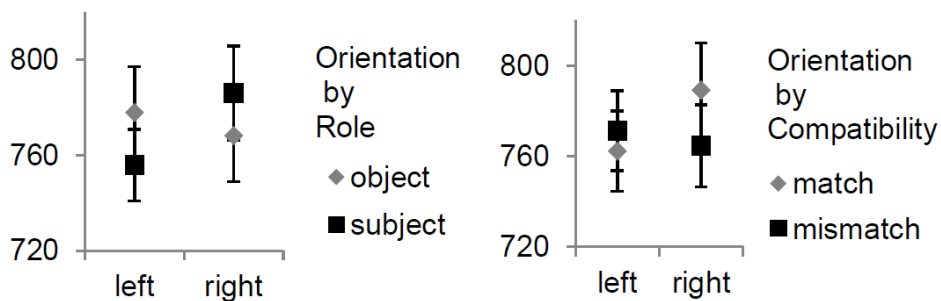
References. [1] Hubbard, E.M., Piazza, M., Pinel, P., & Dehaene, S. (2005). *Nat. Rev. Neuroscience.* [2] Almor, A., Smith, D., Bonilha, L., Fridriksson, J., & Rorden, C. (2007). *NeuroReport.* [3] Konieczny, L., Haser, V., Muller, D., Weldle, H., Wolfer, S., & Hemforth, B. (2010). *23rd CUNY Conf. on Human Sentence Processing.*
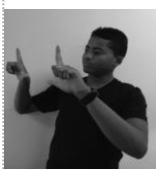
Poster Session 1

Weak and strong definites in sign language

Thaís Sá, Guilherme Lourenço de Souza & Maria Luiza Cunha Lima (Universidade Federal de Minas Gerais)
marialuiza.cunhalima@gmail.com

Definite noun phrases such as "*the TV*" are generally viewed as referring to a uniquely identifiable entity. Carlson and Sussman (2005) proposed that sometimes definite NPs receive a weak reading, in which the uniqueness aspect is not central. In this study we turn to sign languages – specifically Brazilian Sign Language, Libras – to explore the potential of its spatial marking of reference to shed light on the weak and strong distinction. Sign languages rely on space to express reference, thus presenting different morphological encodings for reference when compared to oral languages: after introducing referents, signers assign a definite location for it in the space around their bodies. Whenever they refer back to those referents, signers point to the place to which the referent was assigned. Various authors have proposed that the mode of referent introduction in sign languages plays the same role articles and other determiners play in oral languages (Bahan et al, 1995).

In our first experiment, we elicited the production of strong and weak definites by presenting eight Libras' deaf signers with 36 short videos with target items ("the TV" in the table below, with English translations) signed in the two conditions. Each target item was mentioned twice by different characters. This tested if the two mentions would be interpreted as coreferential - therefore uniquely identifiable and strong - or not. Subjects retold the sentences combining them into one. Results show that when introducing weak definites, signers use the space right in front of their bodies, which we called a neutral space, for both occurrences of the target items. In contrast, in the strong condition, signers introduced the referents either to the left or right hand side of their bodies in the first occurrence and pointed to that determined space in the second occurrence, indicating coreferentiality and therefore uniqueness.

In a comprehension study, we presented deaf Libras signers with short videos in which referents were introduced with either the determined or neutral spaces. Subjects were asked to manipulate a visual display to match pictures of the referents to the characters of the sentences. Subjects tended to use only one picture of the target item for both characters in the strong condition and two pictures of the same item in the weak condition. Results show that Libras' signers reliably use the place of referent introduction to distinguish between uniquely identifiable and non-uniquely identifiable referents. This morphosyntactic distinction supports the interpretation proposed by Carlson and Sussman(2005) between strong and weak interpretations of the definite noun phrases.

WEAK DEFINITES	STRONG DEFINITES
Character 1: I saw the accident on the *TV*. Character 2 : I also saw the accident on the *TV.*	Character 1: I bought the *TV* today. Character 2 : I loved the *TV*.
TV introduced in the neutral space	*TV* introduced in the determined space

References

Carlson, G. and Sussman, R. (2005). Seemingly Indefinite Definites. In S. Kespar and M. Reis (eds.), *Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives*: Mouton de Gruyter.
Bahan, B et al. (1995). Convergent Evidence for the Structure of Determiner Phrases in. *VI Proceedings of the Sixth Annual Meeting of the Formal Linguistics Society (pp. 01-12)*. Bloomington: Indiana University Linguistics Club.

What happened (and what didn't): Discourse constraints on alternative sets

Scott Fraundorf, Aaron Benjamin, & Duane Watson (University of Illinois at Urbana-Champaign)
scottfraundorf@gmail.com

We investigated constraints on the alternative sets generated in comprehending written discourse. Evidence that alternative sets contribute to the long-term representation of discourse comes from recognition memory. [1] tested memory for discourses in which a *context passage* established a set of alternatives (*Jupiter* and *Saturn*, 1a) and a *continuation* mentioned one item from the set (*Jupiter,* 1b). Contrastive prosody on this item facilitated later rejections of a false statement about the alternative (*Saturn)* but not of an unmentioned item (*Neptune).* This pattern suggests the prosody promoted encoding of a specific alternative to the correct item. Although it has been argued that context constrains the set of encoded alternatives [2], it is unclear *which* aspects of context matter. We tested whether mention in the discourse was enough to establish a referent as an alternative, or whether alternative sets were further constrained by the situation model [3] of the discourse. We also tested whether the representation of alternatives, previously observed in spoken discourse, generalized to font emphasis in written discourse, which has been argued to have a contrastive reading [4].

Experiment 1 (*N*=48). Written versions of [1]'s materials were presented for self-paced moving window reading. The critical word in each continuation (*Jupiter* in 1b) was manipulated to either receive font emphasis or not. (Between subjects, emphasis varied between *italics* and CAPITALS; both produced the same memory effect.) After reading all the passages, participants took a true/false memory test (2) containing true probes (*Jupiter*), false probes about the alternative (*Saturn*), and false probes about an unmentioned item (*Neptune*).

We assessed memory performance using multi-level models. Emphasized text facilitated rejection of the alternative (z=2.05, p<.05) but not of the unmentioned item (z=0.11, p=.91), generalizing the representation of alternatives in discourse [1] to emphasis in written text.

Experiment 2 (N=48). In Experiment 2, the initial context included an additional *merely mentioned* item, such as *Neptune* in (3). This item was from the same semantic category as the target, but the discourse established it as a less likely alternative. (The lexical items assigned to the alternative and merely mentioned conditions were counterbalanced across lists.) If prior mention is sufficient to establish an item in the alternative set, emphasis should equally help reject the alternative (*Saturn*) and merely mentioned item (*Neptune*). But if plausibility constrains the alternative set, emphasis should help reject only the plausible alternative.

Emphasis was manipulated using capitals. Emphasis reliably facilitated rejection of the alternative, z = 3.47, p < .001, but not of the merely mentioned item, z = 1.07, p = .28. This effect is unlikely to be an artifact of baseline rejection rates, which were similar across conditions (alternatives: 65%, merely mentioned: 70%).

Discussion. Font emphasis facilitated rejections in memory of a false statement about an alternative to the emphasized item but not rejections of a merely mentioned item. This result indicates that alternative sets encoded in response to emphasis are constrained by the situation model and not only by prior mention. These results also generalize the memory benefits of contrastive information to fonts in written discourse.

1a.   *Expt. 1 context*: Originally, the space probe Cosmo III was designed to fly past Jupiter and Saturn and send photos back from both planets.
1b.   *Continuation*: However, a glitch caused the photos taken of (Jupiter/JUPITER) to be lost.
2.   *Test probe:* Data from (Jupiter/Saturn/Neptune) was lost due to a bug in the space probe.
3a.   *Expt. 2 context:* Originally, the space probe Cosmo III was designed to fly past Jupiter and Saturn and send photos back from both planets. NASA needed this information for a future mission to Neptune.

References. [1] Fraundorf, Watson, & Benjamin (2010). *JML*. [2] Rooth (1992). *Nat'l Language Semantics*. [3] Zwaan & Radvansky (1998). *Psych Bull*. [4] McAteer (1992). *Applied Cog Psych*.

Poster Session 1

Implicit prosody and contextual bias in silent reading

Katherine McCurdy (European Master of Clinical Linguistics), Gerrit Kentner (Goethe University of Frankfurt) &
Shravan Vasishth (University of Potsdam)
vasishth@uni-potsdam.de

The experience of 'hearing' inner speech during silent reading has long been noted, but its nature remains poorly understood. Current eye-movement studies of implicit prosody have uncovered effects at the level of lexical stress, suggesting that implicit meter interacts with syntactic category assignment (Breen & Clifton, 2011).

A recent study (Kentner, 2012) found evidence of implicit prosody guiding parsing in syntactic category assignment of the ambiguous German lexeme *nicht mehr*. In this phrase, *mehr* can be analyzed as part of a temporal adverbial phrase meaning 'not anymore', with main stress on *nicht*, or as the comparative syntactic complement to the verb, which requires main stress on *mehr*. When *nicht mehr* was followed by a verb with stress on the initial syllable (e.g. <u>nach</u>weisen) rather than the medial syllable (e.g. er<u>mit</u>teln), a preference to avoid stress clash led readers to assign *nicht mehr* the temporal reading more often, resulting in evidence of greater processing cost when the sentence disambiguated to the comparative reading. The following example illustrates this ambiguity:

1. Der Polizist sagte, dass man nicht mehr [<u>nach</u>weisen/er<u>mit</u>teln] kann, [wer der Täter war/als die Tatzeit].
2. *The policeman said that one couldn't [prove/determine] [anymore who the culprit was/more than the date of the crime].*

The present eye-tracking study investigates whether contextual bias can modulate the effects of implicit meter. If context shows an early influence on structural processing (Altmann & Steedman, 1988), then the prosodically-induced garden path on *nicht mehr* described above might be reduced or eliminated by a preceding discourse context favoring the comparative reading.

Participants (N=48) read eighteen test sentences containing *nicht mehr*, all of which disambiguated to the comparative reading. Fillers included eighteen sentences in which *nicht mehr* resolved to the temporal reading, ensuring an equal proportion of sentences with comparative and temporal readings. All items were preceded by a context sentence with a bias toward one of the two readings. The biasing effect of context was assessed in a prior rating study. The factors of verb stress (initial vs. medial) and contextual bias (temporal vs. comparative) were crossed in a 2 × 2 Latin Square design and balanced in four lists across test and filler sentences.

In the three-word ambiguous region beginning at *mehr*, main effects of verb stress were dominant, while early effects of context were almost entirely absent. Initial stress on the main verb yielded a significant increase in first-pass regressions on two of the three words, and a significant increase in re-reading probability for all three words. In the two-word disambiguating region, the disambiguating word itself (*als*) showed increased processing difficulty (lower skipping and increased re-reading probability) when verb stress produced a temporal bias, while the word immediately following showed main effects of context in those same measures. Taken together, effects of initial verb stress upon eye movements were swift and pervasive across first-pass and second-pass measures, while effects of context were relatively delayed. These results indicate a strong role for implicit meter in guiding parsing, one that appears insensitive to higher-level constraints.

References. [1] Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, 30(3), 191– 238. [2] Breen, M., & Clifton, C. (2011). Stress matters: Effects of anticipated lexical stress on silent reading. *Journal of Memory and Language,* 64(2), 153–170. [3] Kentner, G. (2012). Linguistic rhythm guides parsing deci- sions in written sentence comprehension. *Cognition*, 123, 1-20.

How focus particles like 'only' hamper the rejection of contrastive alternatives

Nicole Gotzner, Katharina Spalek (Humboldt-Universität zu Berlin), & Isabell Wartenburger (Universität Potsdam)
nicole.gotzner@hu-berlin.de

Listeners are sensitive to contrastive alternatives in online language comprehension. For example, previous research has shown that contrast-evoking prosody leads to the activation of contextual alternatives to an accented word (e.g., Braun & Tagliapietra, 2010). Such alternatives play a crucial role in the definition of certain particles. So-called focus particles like *only* must refer to a contextually-salient set of alternatives (cf. Rooth, 1992) and the presence of a focus particle has been found to facilitate recall of focus alternatives (cf. Gotzner, et al., 2012). The current study used an immediate recognition memory paradigm to test how focus particles affect encoding of alternatives that are (1) explicitly mentioned or (2) implicit in the context.

In the first experiment (n=42), participants were presented with auditory dialogs that introduced a set of elements and mentioned a person and an action (e.g., *In the fruit bowl there are pears, cherries and bananas. I bet Anna ate cherries and bananas.*). In the third sentence, one of the elements was mentioned again and carried intonational focus. The critical manipulation was whether the sentence contained the particle (a) *only*, (b) *even,* or (c) no particle (control) (e.g., *No, she only/even/_ate pears*) while the pitch accent was held constant. Immediately after exposure to the stimuli, the participants had to indicate whether a probe had appeared in the discourse and we measured the time it took (1) to recognize a target that was part of the introduced alternative set (*cherries* in the example), (2) to correctly reject a semantically-related foil (e.g., *apples*) and (3) to reject an unrelated probe (e.g., *socks*). The results showed that the rejection of the foils was slower in the two conditions with focus particles (a,b) relative to the control condition (c) (mixed effects model: *only* vs. no particle $p<.05$; *even* vs. no particle $p<.05$). Crucially, the unrelated probes did not show any differences across particle conditions (a,b,c). The effects of the particles on the recognition of correct alternatives were also not significant indicating that the particles neither facilitated nor hampered immediate recognition of contextually-given alternatives.

The second experiment (n=24) examined the rejection of contrastive alternatives more closely. Typically speakers do not provide a listener with an explicit set of alternatives, rather, this set needs to be reconstructed from the context. Therefore, we reasoned that participants should be able to infer a set of alternatives from the particles' meaning even if the context does not provide the alternatives explicitly. Subjects were either asked to reject a contrastively-associated probe word (nouns that could be substituted for a given focused element, e.g., *dog* vs. *cat*) or a non-contrastive associate (e.g., *dog* vs. *leash*). It was found that the particles *only* and *also* inhibited correct rejections of contrastively-associated probes (mixed effects model: *only* vs. control $p<.05$; *also* vs. control $p<.05$) while they did not affect the rejection of non-contrastive associates. This selective effect suggests that the presence of a particle encourages hearers to entertain alternatives to the expression in focus that can be replaced with it. Taken together, our results indicate that focus particles lead to additional processing costs when participants correctly reject potential alternatives to a focused expression. We assume that the particles encourage a hearer to infer a set of alternatives and this leads to an increased difficulty to decide that a foil had not appeared in the discourse. The study thereby extends previous research on the activation of contextual alternatives to the domain of focus particles. For the first time, we demonstrate that listeners are able to infer a set of alternatives solely based on the semantic definition of a focus particle and that this has an in impact on probe recognition memory.

References

Braun, B., & Tagliapietra, L. (2010). The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes, 25*, 1024-1043.
Gotzner, N., Spalek, K. & Wartenburger, I. (2012, September). *The impact of focus-sensitive particles on memory for focus alternatives*. Poster presented at AMLaP 2012, Riva del Garda, Italy.
Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics, 1*, 75-116.

Poster Session 1

The effect of predictability in elided vs. non-elided constituents

Alex B. Fine & Jeffrey T. Runner (University of Rochester)
abfine@gmail.com

Verb phrase ellipsis (VPE) has received a great deal of attention in both formal linguistic [1-2] and psycholinguistic [3-5] research. The major question from both perspectives concerns the nature of the linguistic representation that must be retrieved or constructed at the ellipsis site. Here, we ask a complementary question: what role, if any, does the contextual predictability of the elided VP play during the comprehension of ellipsis? We know that highly predictable linguistic material is generally easier to process than less predictable material [6]. **Does predictability have the same effect when processing phonologically null material?** If so, this would seem to indicate that (i) the online interpretation of ellipsis is sensitive to detailed lexical and semantic information in the elided VP, and (ii) the notion of predictability relevant to language comprehension may be less strictly form-based than is often tacitly assumed. **Experiment**. To examine the effect of predictability during the comprehension of VPE, we conducted a self-paced reading experiment in which subjects were presented with sentence pairs such as (1)-(4) in which ellipsis and VP predictability in the second conjunct were manipulated. Predictability was operationalized as the subjectively assessed "typicality" of the VP (*cited the book*) given the NP (*The chemist*). 48 subjects read 24 sentences with ellipsis and predictability manipulated within participants and counter-balanced across experimental lists. The effects of ellipsis, predictability, and the interaction between the two on length-adjusted reading times during the underlined region (the auxiliary, the word *too*, and the following two words to allow for spillover effects) were assessed using linear mixed effects regression. **Results**. There was a main effect of ellipsis ($p<.05$): elided VPs were read overall faster than non-elided VPs. This effect held after controlling for the effects of both word length and the number of words per region. The effect of predictability was marginally significant ($p=.056$): VPs that were highly typical or highly predictable were read more quickly than VPs that were lower in predictability, replicating scores of comprehension studies finding effects of contextual predictability [6-8]. Crucially, though, the effect of predictability was qualitatively identical across ellipsis conditions, and the two-way interaction was not significant ($p=.44$). **Discussion**. Our study is the first we are aware of to suggest a global processing advantage for elided compared to non-elided constituents during online language comprehension, and the first to demonstrate an effect of predictability during the comprehension of phonologically null linguistic elements. The results have potential implications for theories of the role of prediction during language processing and for what types of linguistic information must be included in the representations that are accessed in the interpretation of ellipsis.

1. **Ellipsis/High Predictability:** The chemist cited the book in her paper. Later, the physicist <u>did too after seeing</u> its relevance.
2. **No Ellipsis/High Predictability:** The chemist cited the book in her paper. Later, the physicist <u>cited it too after seeing</u> its relevance.
3. **Ellipsis/Low Predictability:** The chemist cited the book in her paper. Later, the cook <u>did too after seeing</u> its relevance.
4. **No Ellipsis/Low Predictability:** The chemist cited the book in her paper. Later, the cook <u>cited it too after seeing</u> its relevance.

References 1. Merchant (2011) **2.** Sag (1976) **3.** Tanenhaus & Carlson (1990) **4.** Arregui et al. (2006). **5.** Kim et al. (2012) **6.** Hale (2001) **7.** MacDonald et al. (1994) **8.** Levy (2008)

Accents and boundaries both affect attachment

Katy Carlson (Morehead State University)
k.carlson@moreheadstate.edu

In a sentence like (1), with a final adverbial which could attach to the higher clause (*claimed last week*) or the lower clause (*arrived last week*), we know that prosodic boundaries affect attachment. All else being equal, a prosodic boundary separating the adverbial from the nearest verb will favor high attachment [1]. In this study, we also manipulated the location of a contrastive pitch accent on the higher or lower verb, on the hypothesis that an accented verb might draw modification by the adverbial. Indeed, both the presence of a prosodic boundary and the location of accent significantly affected the interpretation of the sentences. This finding shows that pitch accents, in addition to their importance for discourse and information structure, can affect the process of generating the basic syntactic structure of a sentence.

1.	a. John CLAIMED that Mary arrived IPh last week.	30% high attachment
	b. John CLAIMED that Mary arrived last week.	18% high attachment
	c. John claimed that Mary ARRIVED IPh last week.	22% high attachment
	d. John claimed that Mary ARRIVED last week.	12% high attachment

In a written questionnaire (N=28), 20 sentences like this had a strong basic preference for low attachment, with 10% of responses showing high attachment. The sentences were then recorded in the four prosodic conditions shown in (1), with the presence of a late prosodic boundary (an IPh boundary) and the location of a L+H* pitch accent varied. The adverbial had a H* accent in all conditions, and the non-focused verb was unaccented or had a non-prominent H*. In an auditory comprehension study, these conditions were played for listeners who then chose between two visually-presented paraphrases, indicating their interpretation of the sentences. As shown by the percentages in (1), there were significantly more high attachments when a prosodic boundary intervened between the lower verb and the adverbial than without the boundary ($F_1(1,57)=13$, $p=0.001$; $F_2(1,19)=20$, $p<0.001$). There were also significantly more high attachments when the higher verb was accented than when the lower verb was accented ($F_1(1,57)=12$, $p=0.001$; $F_2(1,19)=10$, $p<0.01$). There was no significant interaction.

The finding that accents can affect attachment challenges the usual view of the role of pitch accents. They normally indicate emphasis and the given, contrastive, or new status of information within a context; in combination with the syntactic structure, they show the position of focus within a phrase [2]. These issues, while important, are not as integral to the basic structure of a sentence as attachment. Interestingly, previous research has found that some ellipsis structure ambiguities respond to pitch accents but not to prosodic boundaries [3], supporting the idea that accents and boundaries have different domains of application. This work shows the converse, with both types of prosody affecting syntactic structuring. There is then the question of what drives the accent effect on attachment. It may be that the accent increases the simple memorability or salience of a particular verb [4], and thus increases attachment to it; or the focus which the accent provides, and the related increase in semantic processing, may make one site more attractive for attachment [5].

References

[1] Watson, D., & Gibson, E. (2005). Intonational phrasing and constituency in language production and comprehension. *Studia Linguistica*, *59*, 279-300.
[2] Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics, 1,* 75-116.
[3] Carlson, K., Frazier, L., & Clifton, C., Jr. (2009). How prosody constrains comprehension: A limited effect of prosodic packaging. *Lingua*, *119*, 1066-1082.
[4] Lee, E.-K., and Watson, D. G. (2011). Effects of pitch accents in attachment ambiguity resolution. *Language and Cognitive Processes*, 26, 262-297.
[5] Schafer, A., Carter, J., Clifton, C., Jr., & Frazier, L. (1996). Focus in relative clause construal. *Language & Cognitive Processes, 11*, 135-163.

What counts as given?: Deaccenting and givenness effects in spoken comprehension
Eun-Kyung Lee (Harvard Univ.), Tuan Lam (Northwestern Univ.) & Duane Watson (Univ. of Illinois)
eunkyunglee@fas.harvard.edu

Previous work (Arnold, 2008; Dahan et al., 2002) suggests that listeners interpret reduction as a cue to given information. However, what counts as given information is not clear. In this study, we investigate the constraints on which referents can be candidates for reduced referring expressions, focusing specifically on whether referents must be explicitly mentioned and whether they must be the sole referent that is focused.

In Experiment 1, we tested whether listeners' interpretation of givenness relies on explicit linguistic mention or conceptual evocation. In a visual-world eye-tracking experiment, participants heard two instructions to manipulate referents in a visual display (Fig. 1) containing two cohort objects (sandal/sandwich), an object that was semantically related to one of the cohorts (loafer), and an unrelated object (phone). In the first instruction, one of the cohorts was referred to either explicitly (linguistically given) or implicitly (conceptually given). In the linguistically given condition, the cohort was referred to directly (Put the sandal on the triangle.). In the conceptually given condition, the cohort was mentioned using a superordinate term (Put the shoes on the triangles.). In the second instruction, participants were asked to move either the previously mentioned cohort (anaphoric) or the unmentioned cohort (nonanaphoric) and this expression was produced either with an accent or without (Now put the SANDAL/sandal/SANDWICH/sandwich on the star.). Eye fixations at the onset of the target in the second instruction revealed that listeners reliably expected the previously mentioned cohort as the referent of the reduced form (unaccented) when it was linguistically given (sandal), but not when it was conceptually given (shoes).

In Experiment 1, however, the conceptually given condition (shoes) also differed from the linguistically given condition (sandal) in that the cohort shared focus with another referent in the first instruction. In order to investigate whether the absence of the accent effect in the conceptually given condition was due to shared focus, in Experiment 2, the cohort was conceptually given but was the only referent that was focused (the sandal was referred to indirectly: Put the orange object on the triangle., Fig. 2). There was no evidence that listeners considered the previously mentioned cohort as a referent when hearing the reduced form. The combined results from Experiments 1 and 2 suggest that explicit linguistic mention is required for information to be interpreted as given.

In order to further examine whether explicit linguistic mention alone is sufficient for the givenness effect or whether the referent must also be a highly salient one, in Experiment 2, the cohort was explicitly mentioned, but occurred with another referent that shared focus (Put the loafer and the sandal on the triangles.). Again, there were no reliable effects of accenting or reduction.

Taken together, our findings suggest that listeners' on-line interpretation of givenness may be more restrictive than that of speakers. While both linguistic givenness and visual givenness have been shown to lead speakers to reduce words (Kahn & Arnold, 2012), listeners only consider referents that are explicitly mentioned and highly focused as candidates for reduced referring expressions.
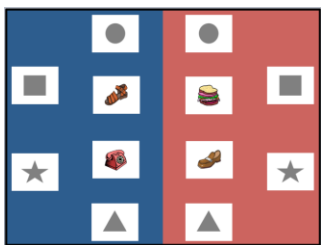

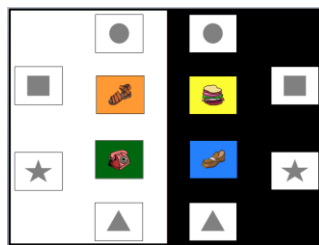**Figure 1.** Example visual display for Expts. 1 & 3

Figure 2. Example visual display for Expt. 2

Effects of distal prosody on perceived word stress and syntactic ambiguity resolution
Nina Gumkowski (Haskins Laboratories) & Mara Breen (Mount Holyoke College)
gumkowski@haskins.yale.edu

Recent studies demonstrate that non-local (i.e., distal) prosodic cues can influence speech perception. Dilley, et al. (2010) demonstrated that distal prosody can influence the perception of word boundaries in ambiguous strings, e.g., listeners perceive *tie/mur/der/bee* as either *timer derby* or *tie murder bee* depending on the prosody of the prior eight syllables. Brown, et al. (2012) demonstrated, in a visual world study, that a similar prosodic manipulation biased listeners to expect targets with specific stress patterns. Sentences like "Heidi sometimes saw that jury" were manipulated to bias perception of the syllable *ju-* (dʒʊə) as strong or weak. Although "that jury" was acoustically identical across conditions, listeners made more early looks to a picture of a jury when the distal prosody predicted stress on *ju-*, but more looks to a giraffe when the distal prosody was consistent with an unstressed *gi-* (dʒʊə).

The current experiment investigated whether stress expectations provided by distal prosody can influence syntactic ambiguity resolution. Thirty-two subjects listened to audio recordings and provided written completions of 24 globally ambiguous sentence fragments (1). The critical conditions (1a,b) ended with stress-alternating homographs like *produce* (noun: *PROduce*, verb: *proDUCE*). For each ambiguous item, there was a corresponding condition (1c,d) with an unambiguous strong-weak noun or weak-strong verb as the final (target) word. The pitch and duration of the first five syllables were resynthesized in 2 ways; in (1a,c), the prosody was intended to bias listeners to perceive stress on the high-pitched syllable of the target, biasing its interpretation as a trochaic noun (Noun-primed prosody); in (1b,d), the prosody was intended to bias listeners to perceive stress on the low-pitched syllable and interpret the target as an iambic verb (Verb-primed prosody). Critically, the homograph itself was acoustically identical across conditions. In order to have a High-Low pitch manipulation of the target words across all conditions and maintain an alternating pitch pattern, the duration of the penultimate word was lengthened in both noun-primed conditions (1a,c).

Completion data demonstrated that, despite a bias to perceive the ambiguous homograph as a noun, the distal prosody biased listeners' interpretations. A binary mixed effects logistic regression on the proportion of noun-consistent completions of the ambiguous conditions (1a,b) revealed significantly more noun-consistent completions to Noun-primed prosody than to Verb-primed prosody, $t = 2.848$, $p < .01$. These results demonstrate not only the generalizability of the distal prosodic effect, but also that perceived stress can influence syntactic ambiguity resolution.

(1)
(a)   Noun-primed prosody, ambiguous   % Noun completions (SE)
        H - L   H   L   H - L   H - L
        Mothers   know   the   g o o d   produce   81.5 (2.0)
(b)   Verb-primed prosody, ambiguous
        L - H   L   H   L   H - L
           Mothers   know   the   good   produce   73.4 (2.3)
   (c)   Noun-primed prosody, unambiguous noun
        H - L   H   L   H - L   H - L
        Mothers   know   the   g o o d   pasta   99.5 (0.3)
(d)   Verb-primed prosody, unambiguous verb
        L - H   L   H   L   H - L
        Mothers   know   the   good   create   6.3 (1.3)

Brown, M., Salverda, A. P., Dilley, L. C., and Tanenhaus, M. K. (2012). *Proceedings of the 34th Annual Conference of the Cognitive Science Society.*
Dilley, L., Mattys, S. L., & Vinke, L. (2010). *Journal of Memory and Language*, *63*, 274–294.

Poster Session 1

A new look at negative sentence verification

Ye Tian, Richard Breheny (University College London), & Heather Ferguson (University of Kent)
ye.tian.09@ucl.ac.uk

In the literature on negation, a widely cited finding for sentence-picture verification studies [1-2] is a polarity/truth-value interaction: while true affirmatives are easier to judge than false affirmatives, participants can take longer for true negatives than false negatives (TA < FA < FN < TN). Accounts for this effect either propose a specific verification procedure for propositional representations [1-2] or a specific mode of representing negative content, which involves first representing the positive argument of negation [3].

We apply insights from dynamic semantics/pragmatics to language processing research to re-interpret this pattern. In current dynamic approaches, interpretation is the update of an information state which includes information bearing on the purpose of the utterance, commonly described in terms of Questions Under Discussion (QUDs) [4]. As demonstrated in [5], when presented with a simple negative sentence without context (for example 1a), participants will probabilistically infer a QUD wherein the positive proposition is at issue, (2a), as part of the comprehension process. In verification studies, the task QUD is (2c/d), so the sentence QUD, (2a), may interfere. In particular its incongruence with the negative items interferes with the verification process, and thus triggers a change in strategy: Initially, participants spontaneously infer what kind of situation supports the negative proposition, and verify such a situation with the picture. In this case, we should observe the same pattern for both positive and negative sentences, i.e. TA < FA < TN < FN. However, this strategy is hard for negatives due to the QUD interference. An alternative strategy is to answer the sentence QUD, (2a), and then, for negative sentences, reverse the polarity of that answer to do the task. This strategy will predict the interaction pattern TA < FA < FN < TN. We note that the interaction pattern is found in classic studies where long training blocks precede a lengthy, repetitive testing phase.

We made two predictions: (i) in the classic paradigm, the RT interaction will only emerge in the late phase, (ii) No strategy change will be observed if we remove or reduce QUD interference in negative sentences (by adding a second object to the picture). One of the two pictured objects satisfied the predicate of the sentence (e.g. a picture of a whole egg alongside cracked sunglasses). Against this picture context, the QUD for a negative sentence is (2b), which is congruent with the negative target sentence (1a). Two groups (one-object or two-object picture) of participants saw 140 sentence (as in 1a,b,c or d) and picture pairs in a verification task. Results: In the one-object group, there was a pattern change: we found a main effect of polarity $(F(1,39)=210.6, p<.001)$ and a main effect of truth value $(F(1,39)=29.7, p<.001)$ in the early phase (first half), but an interaction between polarity and truth value $(F(1,39)=6.7, p=.01)$ in the late phase. Together there is a significant polarity*truth-value*early-late three way interaction $(F(1,39)=14.1, p=.001)$. This training effect is also observed in a longer (320 trials) experiment using the original Clark and Chase items. In the two-item group, no interaction pattern developed. We found main effects of polarity $(F(1,39)=144.5, p<.001)$ and truth value $(F(1,39)=68.7, p<.001)$ throughout the experiment. There was no three-way interaction $(F<1.4)$.

(1) a. The egg is not cracked.   (2) a. Is the egg cracked?

(1) b. The egg is cracked.   (2) b. Which is not cracked?

(1) c. It is the egg that is not cracked. (2)c. Is it true that the egg is cracked?

(1)d. It is the egg that is cracked.   (2) d. Is it true that the egg is not cracked?

References: [1] Carpenter & Just (1975). Psych. Review, 82; [2] Clark & Chase (1972). Cog. Psych. 3; [3] Kaup, Lüdtke & Zwaan, (2005). Proc of CogSci., 27; [4] Ginzburg (2012). The Interactive Stance. OUP; [5] Tian, Breheny & Ferguson (2010). QJEP, 63.

The hypothetical property of "if"-statements: A visual-world paradigm eye-tracking study
Likan Zhan, Stephen Crain, Peng Zhou; (Macquarie University)
likan.zhan@mq.edu.au (Likan Zhan)

This study used eye-tracking in the visual world paradigm to test the hypothesis that "if p" is assertable in a situation where the truth of "p" is not acknowledged. In asserting an "if p, then q", speakers (hypothetically) add the content of "p" to their current stock of knowledge and make inferences about the truth of "q". Statements with "because" differ from conditional statements in that "because p" is only assertable in situations where "p" is known to be true. In the example (Fig. 1, right panel), the princess has eaten the apple, but has not eaten the banana. If the hypothesis is correct, then the object in the test image that satisfies the "if"-statement, "if the princess eats that __", is the banana, not the apple. By contrast, the only object that satisfies the corresponding "because"-statement, "because the princess ate that __", is the apple, not the banana. We recorded participants' fixations on test images (e.g., Fig.1, right panel) while they listened to test audios in Mandarin Chinese (Fig. 1, left panel). The experiment intended to see if the semantic contributions of the two expressions had sufficient psychological reality so as to guide participants in fixating more on the banana (area CF, Fig. 1, right panel) in the "if"-statement, and more on the apple (area FA, Fig. 1, right panel) in the "because"-statement. We tested 40 Mandarin-speaking adults on 108 test trials (18 images × 3 connectives × 2 objects). For the analysis, the temporal regions of interest began at the onset of the sentence, and partitioned the sentence in 39 temporal bins of 100ms each, ending at the onset of the mentioned object, "banana" or "apple," at 3.9sec. The dependent variable was the proportion of fixations to an object (banana vs. apple) in a bin of interest divided by all fixations recorded in that bin (see Fig. 2). The data were then fitted with GLMMs (R: proportion ~ Bin + I*Bin^2 + I*Bin^3 + connective + (1 | subject) + (1 | trial)). The results confirmed our hypothesis (Fig. 2). On the banana, more fixations in "if"-statement (p<.001), while less in "because"-statement (p<.01) were found, compared to the baseline "when"-statement. To the apple, by contrast, more eye-movements in "because"-statement (p<.01) and less in "if"-statement (p<.001) were found, compared to the baseline. Further analyses in each bin (R: proportion ~ connective + (1 | subject) + (1 | trial)) suggest that the effects mainly occurred after the onset of the verb (p<.01) (Fig. 2., dotted rectangles).
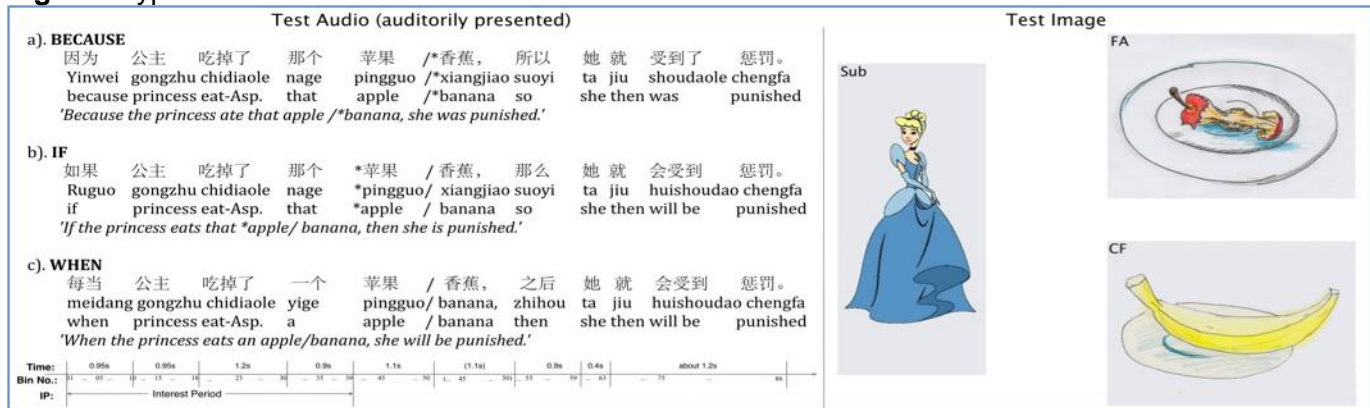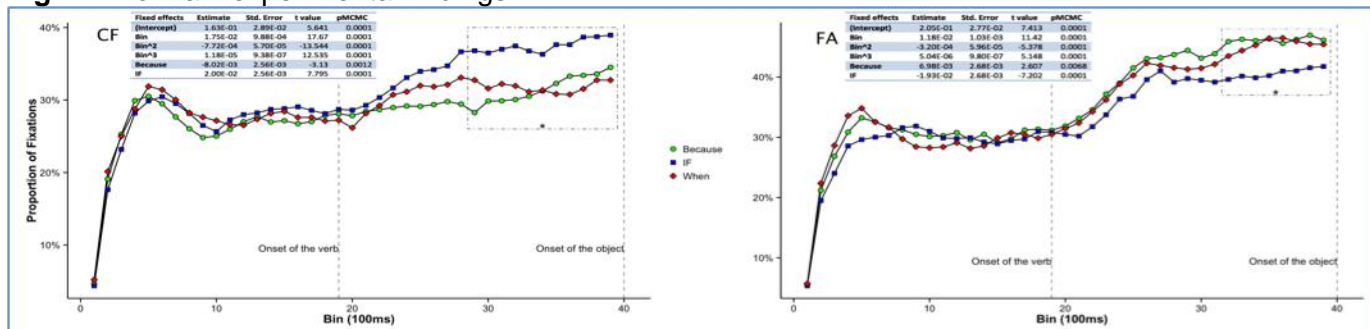
Fig. 1. A typical trial.

Fig. 2. The main experimental findings.

Implicatures in uncooperative contexts: Evidence from a visual world paradigm

Anna Pryslopska (SFB833, University of Tübingen)
anna.pryslopska@student.uni-tuebingen.de

Scalar implicatures are a hot topic among linguists and there is much controversy as to how they are computed. Neo-Griceans [4] argue that semantic and pragmatic meanings are processed in two hierarchically ordered stages, causing implicatures to be delayed in comparison to purely semantic quantifiers like *all*. Defaultists [3] claim that listeners arrive at the pragmatic meaning automatically and not later than e.g. for *all*. Relevance theorists are divided in their assessment; some [2] claim implicatures can be computed immediately and effortlessly, while others [1] consider it a time and resource consuming process. The present study addresses this debate in a visual world experiment which is partially a replication of Grodner et al., but also investigates whether implicatures are computed when a prerequisite for communication in Gricean terms – cooperativity – is lacking. Will implicatures be computed immediately (defaultists), will they be delayed (relevance theorists) or even absent (Neo-Griceans)?

A pretest aimed at confirming that the German quantifier *einige* (*some*) carries an implicature, as it does not fully equal *some*. 64 fillers and 16 items in 4 conditions were distributed over 4 lists. The conditions were pairs of pictures (7x7 matrix with 4 colors) and quantified sentences (eg. "In my bag are <u>some</u> red pebbles."): quantifier *all* was paired with a picture where all pebbles were of the relevant color and *none* with a corresponding none-picture; *some* was paired with two pictures, some and all. A condition of false pairs was added for control. 32 German native speakers judged the pairs' acceptability on a 7-point scale with the option to refute the pair as unacceptable. *Some+some* conditions judged better that *some+all* ($F_1(1,29)=132.4$). False controls were rejected 92.4%, items <11% of the time.

Materials in the eye-tracking experiment were almost identical to the pretest. The factors were quantifier type, picture (both: *some*, *all*, *none*), and speaker (coop., uncoop.). 2 randomized experimental lists were counterbalanced for speaker gender. 112 fillers and 24 items in 6 conditions: quantifier × speaker paired with the three pictures. The targets for *all* and *none* for both speakers were the corresponding pictures, whereas for *some* the target was *some* for the coop. and *all* for the uncoop. speakers. 22 German native speakers were played a card game with a friendly and opposing confederate where they were forced to make and respond to utterances resembling the items. This ensured that the participants were accustomed to speaker types. In the following eye-tracking part, they were asked to view 3 randomly ordered pictures in a visual world and listen to an utterance made by one of the speakers. The task was to choose the picture that best fits the utterance. With *some* utterances, all-picture for *some* quantifier was chosen significantly more often when the speaker was uncooperative (z=0.0003). A GLMM analysis revealed that for the cooperative speaker, the pragmatic target of *some* was fixated as early as the target of *all*. In uncooperative cases there was a substantial temporal delay between *some* and *all* quantifiers: picture (*some*, *all*) × quantifier (*some*, *all*) × bins (5-10, 11-20) × cooperative (yes, no), z=3.4.

To sum up, in standard, cooperative contexts implicature computation was fast and automatic. However, in uncooperative cases the implicature was computed lateer due to the cancellation and re-computation processes. The results replicate Grodner et al. They are consistent with both the defaultist and the relevance theoretic view [2].

References

[1] Bott, L. and I. Noveck (2004). "Some utterances are underinformative: the onset and time course of scalar inferences." *Journal of Memory and Language* 51(3), pp. 437-457.
[2] Breheny, R., N. Katsos and J. Williams (2006). "Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences." *Cognition* 100(3), pp. 434-63.
[3] Grodner, D.J. et al. (2010). ""Some," and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment". *Cognition* 116(1), pp. 42-65.
[4] Levinson, S. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press.

Focus inhibits free associates

Mary Byram Washburn, Elsi Kaiser, Maria Luisa Zubizarreta (University of Southern California)
byram@usc.edu

Contrastive focus is frequently analyzed as introducing alternatives to a contrastively focused constituent (c.f. Rooth 1992), such that a sentence like 'JANE loves Mark' evokes alternatives to Jane, the contrastively focused constituent (e.g., Pam, Amy, ex.1).

1) **Jane** loves Mark: $\exists x \exists y . x \in C \, \& \, y = Mark \, \& \, loves(x, y) \, \& \, C = \{Jane, Pam, Amy\}$

Past experimental work has shown that contrastively focusing a constituent makes it easier to access its alternatives. Norris et al (2006), Braun & Tagliapietra (2009), and Byram Washburn et al. (2012), in priming studies, found that alternatives to a contrastively focused prime are accessed faster than non-alternatives. The current work explores whether this facilitation depends on the semantic nature of the alternatives--in particular, whether the alternatives are (a) *semantic associates* of the target word (as identified by free-association norms, Nelson et al., 1998) or (b) *contextual associates* evoked by the sentential context (but crucially not semantically associated with the target word). To preview our results: Using mouse-tracking methodology which provides fine-grained data about the alternatives that people are considering (Freeman & Ambady, 2010), we find that semantic alternatives–but not contextual associates–are inhibited.

Method. Participants (n=24) saw two words in the top corners of the computer screen, and *heard* a sentence (e.g. "Roberta took an art class so she painted a cliff."). The final word was the target. Participants were instructed to use the mouse to click on the word that best fit the end of the sentence. **Conditions**: We manipulated the prosody of the auditorily-presented sentence and the nature of the visually-displayed words shown on the computer screen: The target word was prosodically focused (contrastive L+H* accent) or unfocused. On the screen, participants saw (a) a semantic associate (*ledge*) and a nonsense word (*twarked*), or (b) a contextual associate (*portrait*) and a nonsense word (*twarked*), OR (c) a semantic associate and a contextual associate. The semantic associates were selected using the University of South Florida free association norms, and had an average association strength of .21 to the target. The contextual associated were selected by means of a fill-in-the-gap norming study so they were contextually predictable. Crucially, they were not free associates of the target.

Results. Mouse tracking measures the difference between the actual mouse trajectory and an ideal trajectory (straight line to the chosen word). Deviations towards the distractor indicate that the participant is considering the distractor (cf. fixations in eyetracking, Spivey et al., 2005). We found that when participants saw **a semantic associate and a nonsense word,** they clicked on the semantic associate, as expected, regardless of whether the prime was focused. Crucially, mouse trajectories deviated away from the semantic associate *more* when the prime was CONTRASTIVELY focused than when it is was unfocused (p's<.05 for maximum deviation). This suggests that contrast suppresses activation of semantic associates. When participants saw **a contextual associate and a nonsense word**, they ultimately clicked on the contextual associate, regardless of focus. Mouse trajectories in this condition show *no indication of an effect* of prosodic focus influencing the processing of contextual associates. This suggests that the set of alternatives is composed only of semantic associates of the focused constituent. Interestingly, though, when the participants saw a **semantic associate and a contextual associate**, they generally favored the semantic associate and tended to choose it even more when the prime was focused. This suggests that alternatives might be initially inhibited, but eventually receive additional activation.

Summary. As a whole, our results suggest that focus causes the set of alternatives to be, at least initially, inhibited, and that the set of alternatives evoked by contrastive focus is composed only of the semantic associates of the focused constituent.

Incremental computation of scalar implicatures: An ERP study

Les Sikos, Sam Tomlinson, Hilary Traut & Daniel Grodner (Swarthmore College)
lsikos1@swarthmore.edu

Recent evidence on the time course of processing Scalar Implicatures (SIs) is mixed. Some results suggest that SIs are processed immediately at a scalar expression (Breheny et al. 2012; Grodner et al. 2010), while others suggest that SIs are delayed relative to their literal meanings (Huang & Snedeker 2009; Bott & Noveck 2004). Here we investigated the incremental processing of SIs in two experiments using event-related potentials (ERPs). Previous ERP work has demonstrated that underinformative clauses (*Some people have lungs*) elicit processing difficulty (reflected in an N400) compared to informative clauses (*Some people have pets*), though only for pragmatically skilled participants (Nieuwland et al. 2010). The goal of the present study was to compare scalar vs. non-scalar quantifiers (Exp 1: *some* vs. *all*; Exp 2: *many* vs. *all*) within three levels of informativity (Table 1). This design allowed us to compare brain responses at the quantifier, sentence-final word, and verification. Individual differences in pragmatic skills were also assessed via the Autistic Spectrum Quotient.

Methods. Participants (N_1=48; N_2=30) read and responded to 216 statements (36 per condition) from a naïve speaker in a simulated dialogue. Following Bott & Noveck (2004), participants were trained to interpret scalar quantifiers as implying *not all*. Each trial was followed by a delayed (500ms) sentence-verification task.

Results and Conclusions. Exp 1: ERPs at the scalar quantifier *some* diverged negatively from *all* in an early window (230-440 ms) post word onset. We propose that this "scalar effect" may reflect anticipatory processes related to scalar quantifiers providing a functional signal that more complex conceptual integration is forthcoming. Consistent with this hypothesis, at sentence-final target words, scalar-UI elicited more left-frontal negativity than non-scalar-UI 300-400 ms post word onset. Finally, verification elicited differences 1500-2500 ms post target word onset, which patterned with difficulty (accuracy; response time) rather than with quantifier type. These results suggest that computation of a SI begins at the scalar quantifier and influences expectations for upcoming words, leading to semantic processing costs for words that deviate from expectations.

ERPs for high and low pragmatic-skill groups (median split) exhibited largely the same patterns. However, only high-skill participants showed a larger N400 for scalar-INF vs. non-scalar-INF (300-500 ms). This extends prior findings that individuals with greater communication skills make immediate use of pragmatic information.

Exp 2 replaced the scalar quantifier *some* with *many* and replicated several key findings from Exp 1: *many* elicited an early left frontal negativity relative to *all,* demonstrating that the scalar effect is not idiosyncratic to *some*. Further, sentence-final words in scalar-UI sentences were more negative than in non-scalar-UI. However, both of these effects were less robust and more fleeting than Exp 1. High-skill participants again showed significantly greater processing activity for scalar-INF.

This work supports the view that SI calculation begins immediately at the scalar expression. It also extends previous results to the weaker scalar quantifier *many*.

Condition	Example stimulus Item	Veracity
Scalar-INF	I believe that <u>some/many</u> people have <u>pets</u>.	T
Non-scalar-INF	I believe that <u>all</u> people have <u>pets</u>.	F
Scalar-UI	I believe that <u>some/many</u> people have <u>lungs.</u>	F
Non-scalar-UI	I believe that <u>all</u> people have <u>lungs.</u>	T
Scalar-False	I believe that <u>some/many</u> people have <u>planets</u>.	F
Non-scalar-False	I believe that <u>all</u> people have <u>planets.</u>	F

Stress position congruency hinders word production: Evidence from the picture-word interference paradigm

Claudio Mulatti (University of Padua), Simone Sulpizio & Remo Job (University of Trento)
simone.sulpizio@unitn.it

In any polysyllabic language with no fixed stress position, a core issue for speech production is how speakers retrieve and assign stress to words. The model of word production of Levelt, Roelofs, and Meyer (1999) posits that, during word phonological encoding, the retrieval of the word metrical representation occurs in parallel with the retrieval of the segmental information and, crucially, that these two processes take the same amount of time. A consequence of such architecture is that priming either of the two processes should not affect total processing time, since the system has to wait for the slower, unprimed process to complete its computations (cf. Roelofs & Meyer, 1998). However, if one of the two processes (e.g. the retrieval of the metrical representation) is slowed down, then this slowing should affect total processing time, since the system has to wait for this slower process to complete its computation. In the experiment here reported we tested the latter prediction using Italian, a language in which stress is lexically based and has to be retrieved from memory (with an exception only; see Sulpizio, Job, & Burani 2012).

Participants performed a picture-word interference task: to-be-named target pictures were presented along with to-be-ignored, written distractor words superimposed on the target. Targets and distractors had either the stress in the same position (congruent condition, e.g., pa.VO.ne 'peacock' – su.SI.na 'plum') or in a different position (incongruent condition, e.g., pa.VO.ne 'peacock' – MO.du.lo 'form'). Levelt at al.'s (1999) model postulates that units representing a particular metrical structure are shared among words with that particular metrical structure, thus co-activating the same metrical representation during processing. This being the case, the production rule that selects the metrical node in the congruent condition deals with a unit that is connected to both the distractor – and that therefore should not be selected – and the target – and that therefore should be selected. The system needs to accumulate enough information for the appropriate decision to be made and, hence, target naming RTs are predicted to be slower in the stress-congruent condition than in the stress-incongruent condition. Results showed this to be the case, with participants taking 41 ms longer to name target pictures in the congruent (same stress) than in the incongruent (different stress) condition. Since in order to comply with the instructions of a PWI task (i.e. name the target and ignore the distractor) the distractor has to be blocked (Roelofs, Piai, & Schriefers, 2011) we propose that the production rule waits until – after distractor blocking – the activation of the distractor's morpheme unit drops below a given threshold so that the production rule considers it inactive: at this point, only the target nodes are active and the production rule can safely select the activated metrical structure unit. The stress-congruency interference effect thus suggests that metrical information contributes to the organization of the lexicon and may act, at least within the production system, as a clustering parameter of the lexical units.

References

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1-38.

Roelofs, A., & Meyer, A.S. (1998). Metrical structure in planning the production of spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 922-939.

Roelofs, A. Piai, V., & Schriefers, H. (2011). Selective attention and distractor frequency in naming performance: Comment on Dhooge and Hartsuiker (2010). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 1032-1038.

Sulpizio, S., Job, R., & Burani, C. (2012). Priming lexical stress in reading Italian aloud. *Language and Cognitive Processes, 27*, 808-820.

Lexical differentiation in language production and comprehension
Si On Yoon, Sarah Brown-Schmidt (University of Illinois at Urbana-Champaign)
syoon10@illinois.edu

We explored the breadth of the relevant discourse context in production and comprehension. Production research suggests the relevant context is quite broad, based on findings that speakers distinguish new discourse referents from non-present referents discussed in a previous context. For example, after referring to "*the tree*", a different tree is more likely to be called ("*the tree without leaves*") even though only a single tree is currently visible (Van der Wege, 2009).

E1 was designed to replicate this "lexical differentiation" effect. On each trial, a speaker (S, *n*=16) and experimenter (E) viewed four pictures on screens (e.g., shirt, glasses, shoe, ax); the target was highlighted on S's screen. S instructed E to click on the target. In the <u>contrast</u> (C) condition, S described a "contrast" item (e.g., checkered shirt) in entrainment and then the target (e.g., striped shirt) at test. In the <u>non-contrast</u> (NC) condition, S described an unrelated item (e.g., towel) in entrainment, and the target (striped shirt) at test. <u>*E1 Results*</u>. In the NC condition, on test trials Ss referred to the target with a bare noun phrase (89.9%) and a modified noun phrase (10.1%). In the C condition, Ss modified 17.6% and used bare noun phrases only 66%, a significant difference (z=2.81, p<.01). Moreover, these bare noun phrases in the C condition were often distinct from the ones used to refer to the contrast during entrainment (16.4%, e.g., *knife, sword*). Thus, in the C condition, Ss differentiated the target from the contrast through both modifiers and subordinate-level nouns.

E2-3 examined whether *listeners* (L) show differentiation effects by monitoring L's eye movements during the same task. In E2 (*n*=32), two factors were manipulated: (1) Whether the basic level object term for the target (e.g., *shirt*) was previously used to refer to a different token from the same category during entrainment (<u>historical</u> condition, e.g., *Click on <u>the shirt</u>*), or not (<u>ahistorical</u> condition, e.g., *Click on the bottom right object*). (2) Whether the test instruction differentiated the novel target object from the previously-experienced contrast (e.g., *Click on <u>the striped shirt</u>*), or not (e.g., *Click on <u>the shirt</u>*). Like E1, entrainment trials were repeated 6x and there were 1-10 trials (M=6.1) between the last entrainment and test trial. <u>*E2 Results*</u>: Analysis of L's target fixations showed non-modified nouns were interpreted faster in the historical context, whereas modified nouns were interpreted equally rapidly, regardless of the historical context. The results offered no evidence that listeners generate lexical differentiation expectations when interpreting references to a novel exemplar from a previously referenced category term (t=1.362, p>.05). However, listeners might have *expected* lexically differentiated terms, but were unable to accurately predict the specific form of the differentiated term.

E3 used subordinate-level nouns to test for differentiation in comprehension. We hypothesized that subordinate nouns may be more predictable, increasing the likelihood that Ls could make a useful differentiation-based prediction. During entrainment, L (*n*=32) interpreted E's reference to pictures using basic level nouns such as *the dog* (as referring to, e.g., a German Shepard). Test scenes had 4 new objects including a second exemplar from the same object category (e.g., a <u>poodle</u>), a subordinate-level cohort competitor (<u>pool</u>—competes with "*poodle*"), a basic-level cohort competitor (<u>dock</u>—competes with "*dog*"), and an unrelated item (racquet). A cohort-competitor paradigm was used and two factors were manipulated: (1) Entrainment type (historical or ahistorical); (2) Target type on test trial (subordinate level object [e.g., "*poodle*"] or basic level object that was a cohort-competitor with a previously-entrained basic-level noun [e.g., "*dock*"]). Analysis of L's preference to fixate subordinate targets vs. competitors ("*poodle*" vs. pool), and to fixate basic targets vs. competitors ("*dock*" vs. dog (poodle)) showed no evidence of expecting subordinate-level expressions (ps>.05).

The findings suggest that the breadth of relevant discourse context differs across production and comprehension. Ss show more sensitivity to things they have said before, due to better knowledge of the relevant context, whereas Ls have the task of inferring what the S believes is the relevant context; this inferential process may be more error-prone.

Does message similarity facilitate sentence formulation?

Agnieszka E. Konopka (MPI for Psycholinguistics), Stefanie E. Kuchinsky (Medical University of South Carolina), & Antje S. Meyer (MPI for Psycholinguistics)
agnieszka.konopka@mpi.nl

The generation of an utterance begins with event apprehension and continues with sequential linguistic encoding of all message elements [2]. The timecourse of formulation, however, is relatively flexible and varies with the ease of structural encoding [3]. While previous work has shown that syntactic structure may be primed independently of thematic roles across sentences [1], here we tested whether exposure to conceptually similar events interacts with structural processes to facilitate the mapping of a message onto a sentence.

Young adult native Dutch speakers (*n*=41) performed an eye-tracked picture description task with 36 prime-target picture pairs embedded in a list of 160 filler pictures. On target trials, participants described pictures of transitive events (e.g., a dog chasing a mailman). On prime trials, they saw pictures where the action was related or unrelated to the action in the targets and heard a recorded active or passive description (The paparazzi is *following / photographing* the queen; The queen is being *followed / photographed* by the paparazzi) that they had to repeat out loud. The similarity of prime and target events at the conceptual level and lexical level was confirmed with norming (perceptual similarity of related and unrelated picture pairs was held constant). The structural manipulation served to test the joint effects of similarity in conceptual structure and linguistic structure on sentence formulation. We predicted effects of the structural primes (a) on the form of target descriptions particularly when target events were similar to prime events, and (b) on eye movements (fixations to the agent and patient), (c) even in the absence of behavioral structural priming.

As expected, (a) the structural primes influenced the form of target sentences (structural priming), and priming effects were larger between primes and targets with related than unrelated verbs (the semantic boost in structural priming). Timecourse analyses of the eye movements (performed with quasi-logistic regressions and growth curve analyses; GCA) tested at what stage event similarity influenced formulation.

Event Similarity: Event similarity did not influence early eye-movements (i.e., event apprehension, 0-400ms after picture onset), but predicted when speakers deployed attention to the subject character after 400ms. In active sentences, (b) speakers looked briefly at the patient between 400 and 700 ms before re-directing their gaze to the agent; this shift of gaze was less pronounced after related than unrelated prime events, suggesting that speakers found it easier to continue encoding the sentence when describing similar events.

Event Similarity and Structural Priming: After 700 ms, shifts of gaze to the agent in active sentences occurred more quickly after active than passive primes with related verbs, indicating earlier encoding of this character when formulation was supported by a recently used linguistic structure. When producing the dispreferred passive structure, speakers also directed more looks to the patient after passive primes than active primes (GCA), particularly when prime events were related to the targets. Event similarity and Prime structure impacted formulation even when speakers were not behaviorally primed (e.g., produced an active sentence following a passive prime): (c) fixations to the patient tended to be more peaked and shifted earlier after passive primes, and again this effect was larger after exposure to related prime events.

Thus eye movements revealed that message similarity and structural priming can affect formulation for preferred and dispreferred syntactic structures alike and even in the absence of behavioral priming. These effects were observed throughout the linguistic formulation process (from 400 ms onwards), suggesting a tight link between the ease of message formulation and the mapping of this message onto a linguistic structure.

References
[1] Chang, Bock, & Goldberg (2003), *Cognition, 90*, 29-49.
[2] Griffin & Bock (2000), *Psych.Science, 11*, 274-279.
[3] Konopka & Meyer (2011), *AMLaP,* Paris, France.

Incremental planning of complex noun phrases in sentence production

Maureen Gillespie (University of New Hampshire), Victor S. Ferreira (University of California San Diego) & T. Florian Jaeger (University of Rochester)
mtc2@unh.edu

Formulating sentences requires planning of lexical and structural content, and currently there is little consensus about the scope of advance planning in language production. There is evidence that speakers use an extremely narrow scope of planning and plan utterances nearly word-by-word (e.g., Brown-Schmidt & Konopka, 2008; Griffin, 2001; Wheeldon & Lahiri, 2002); however, there is also evidence that speakers plan aspects of full phrases, and even simple sentences, in advance of speaking (e.g., Smith & Wheeldon, 1999; Wagner et al., 2010).

Nearly all scope of planning research examines planning that occurs prior to speech onset, but it is clear that at least some amount of planning occurs as speech is produced (Griffin, 2003). The scope of lexical/grammatical planning of non-subject noun phrases (NPs) was examined in a video description task that elicited semi-spontaneous speech. Sentences took the form "*Bob moved the (big) chair (in the circle) above the dog.*" The critical NP (*chair*) could require prenominal or postnominal modification due to a co-present contrast. Speech onset times (SOTs) and the durations of the first three words which were identical across modifier conditions (Subject, Verb, Object Determiner) were measured. Additionally, durations of the head noun (*chair*) and preposition (*above/below*) appearing after the object NP were also measured to assess whether high planning demands can 'spill over' and affect speech rates at later points in time.

Mixed effect models predicting log duration were used. There were no effects on SOTs. Subject (*Bob*) durations were longer for prenominally modified NPs compared to bare NPs, with the postnominally modified NPs and bare NPs showing equal Subject durations. The Verb (*moved*) and Object Determiner (*the*) durations were longer for postnominally modified NPs compared to bare NPs, with the prenominally modified NPs and bare NPs showing statistically indistinguishable Verb and Object Determiner durations.   There was no effect of NP modification on the duration of the head noun of the object NP nor the preposition following the object NP.

These results suggest that speakers dynamically adjust their speech rate based on planning demands (Fox Tree & Clark, 1997; Griffin, 2003). In one sense, planning of NPs appears to be incremental: The order in which the modifier occurs in the utterance affects when durational lengthening is observed. However, some aspects of planning do not seem to be limited in scope to one syllable, word, prosodic word, or even syntactic phrase. At least some advance   grammatical planning of the object NP takes place during the first few hundred milliseconds of articulation. Under the assumption of cascading planning (Griffin, 2003), the planning scope in production may not be a fixed structural unit; instead, preparation of a particular element (e.g., prenominal vs. postnominal modification) is initiated during a time window that minimizes the probability of becoming disfluent or suspending speech later in the utterance.

References
Brown-Schmidt & Konopka (2008). *Cognition, 109,* 274-280.
Fox Tree & Clark (1997). *Cognition, 62,* 151-167.
Griffin (2001). *Cognition, 82,* B1-B14.
Griffin (2003). *Psychonomic Bulletin & Review, 10,* 603-609.
Smith & Wheeldon (1999). *Cognition, 73,* 205-246.
Wagner et al. (2010). *JEP:LMC, 36,* 423-440.
Wheeldon & Lahiri (2002). *Cognition, 85,* B31-B41.

Modeling word duration in language production

Andrés Buxó-Lugo, Dominique Simmons, Duane Watson (University of Illinois at Urbana-Champaign)
buxo2@illinois.edu

It is a well-known phenomenon that speakers lengthen words that are new, informative, or unpredictable in a conversation while shortening words that are given, predictable or non-informative (e.g. Aylett & Turk, 2004; Bell et al., 2009; Fowler & Housum, 1987; Jurafsky, 2001; and many others). A puzzle for psycholinguists is understanding why. One explanation of these duration effects is that the duration of words partly reflects the complexity of underlying production processes. Speakers produce words that are new or informative with longer duration because those words are actually more difficult to say. Words that are more activated require less articulation time (reduction) while words that are less activated require more articulation time (lengthening). However, if lengthening is linked to planning difficulty, what benefit could a speaker derive from lengthening a new word once articulation has already begun?

In this paper we argue that the serial nature of phonological encoding may lead to reduction and lengthening. In current models of word production (Sevald & Dell,1993; O'Seaghda & Marin, 2000), phonological selection is a serial process, and as such, lengthening might occur at points of complexity in phonological encoding in order to provide sufficient time for encoding to take place. Thus, the strategy we use is to first understand whether a serial encoding model predicts complexity at varying points within a word and across words. Then we test to see whether English speakers' durational choices match predicted points of complexity by the model.

Rather than model phonological encoding we modeled encoding of groups of phonemes that corresponded to morphemes in compounds. This was done to simplify both training of the model as well as measuring durational changes in human productions. We used a simple recurrent network (SRN) inspired by Dell et al. (1993). Two models were trained to produce two-word vocabularies. One model was trained to produce two compounds that overlapped in their initial morpheme (e.g. *layover, layout*). The other model was trained to produce compounds that overlapped in their final morpheme (e.g. *outlay, overlay*). The input to the model was activation of one of two nodes corresponding to the lemma for the target word. The output of both models was a predicted component morpheme (lay, out, over, or a word boundary). On each cycle, the target lemma input node was activated and the model was required to produce the correct output morphemes in sequence. At test, the summed-squared error for the output nodes was higher for SRNs that were trained on word pairs with initial overlap than those trained on words with final overlap. This is consistent with previous interference effects found in the literature (Sevald & Dell, 1993; Jaeger et al. in press). Critically, in both models, the summed-squared error was higher for the non-overlapping morpheme than for the overlapping morpheme (e.g. *over* and *out* in *layover layout*).

To explore whether listeners lengthened words at points predicted by the model, we used a paradigm developed by Sevald & Dell (1993). Fifteen participants produced compounds in alternation as quickly as possible for eight seconds. The compounds either overlapped in their initial morpheme (e.g. *layover*, *layout*) or final morpheme (e.g. *overlay*, *outlay*). The duration of the component morphemes in critical conditions was compared to baseline conditions that did not consist of overlapping compounds (e.g. *layover*, *handout*). We found that points of complexity predicted by the SRNs corresponded with lengthening in participants' productions. Overall, words that overlapped initially were produced with longer duration than words that overlapped finally. Furthermore, slow-downs were significantly greater for the non-overlapping morpheme than for the overlapping morpheme. These data suggest that a serial retrieval process like phonological encoding can result in complexity that is linked to similarity across words. This reveals itself through SRN error and lengthening in human production at similar points. This suggests that, more generally, reduction and lengthening of a word may be linked to the process of phonological encoding.

Poster Session 1

How do speakers think for speaking in a VOS language?

Takuya Kubo[1], Manami Sato[1], Hajime Ono[2] & Hiromu Sakai[1] ([1]Hiroshima University, [2]Kinki University)
takuyak0625@gmail.com

Introduction. According to the 'thinking for speaking' hypothesis, speakers accommodate their thoughts to the demands of linguistic encoding when speaking [1]. To put this into the recent model of language production [2, 3], how people encode message (message encoding) can be influenced by the language they speak (grammatical encoding). Recent studies on nonverbal event description (i.e. depictive gestures) argue that gestural orders reflect a sequencing for representing events at the cognitive level. Although tested languages were limited in SO languages where Subject (Actor) precedes Object (Patient), the Actor-Patient order was strongly preferred in nonverbal as well as in verbal event descriptions [4, 5], suggesting that message planning is closely linked to grammatical encoding. On the other hand, this raises a question of whether speakers of OS languages in which Patient precedes Actor accommodate their thoughts to the Patient-Actor word order. To investigate this issue, we conducted an experiment consisting of verbal and nonverbal event descriptions in Kaqchikel, a Mayan language spoken in Guatemala, whose basic word order is VOS [6].

Experiment. Thirty-one native speakers of Kaqchikel participated in the experiment, involving verbal as well as nonverbal event descriptions. Since Kaqchikel allows SVO order as well as VOS order, we first conducted a verbal description task to investigate the number of VOS sentences they produced. About 6 months later, the nonverbal description task was conducted with the same participants. We predict that if message encoding is influenced by the way in which one speaks, there should be a correlation between word order and gestural order. That is, the more participants produce sentences in the Patient-Subject order, the more they should generate gestures in the Patient-Agent order.

Method. In the verbal description task, participants described 24 line-drawn pictures depicting simple transitive events, in a random order intermixed with 18 filler items. In the nonverbal description task, we selected 18 line-drawn target pictures that were appropriate for the gesture task. Participants described the pictures using gestures to a Kaqchikel collaborator who pretended not to have seen the pictures before.

Results. Large individual differences were observed in the production rate of the Object-Subject word order (ranging from 0-86.4%, average of 18.5%) while they were less obvious in the gestural rate of the Patient-Actor order (ranging from 0-40, average of 9.6%) (Figure1). Analysis using the Pearson's correlation coefficient indicated that the relationship between the rate of Object-Subject sentences and that of Patient-Actor gestures was not significant (r=0.297, p>0.1).

Discussion. First, lack of correlation between the frequency of VOS verbal production and the ordering preference of Patient-Actor gesture production suggests that the language one speaks does not necessarily influence the way one cognitively represents the event. Second, the natural order of event representations is Actor-Patient not only for speakers of SO languages [4,5], but also for speakers of OS languages. Third, our results suggest that the message encoding processes in OS languages may not be distinct from those in SO languages. In sum, the Patient-Actor word order in OS languages may not be generated in the message planning level, but is determined in the grammatical encoding level.
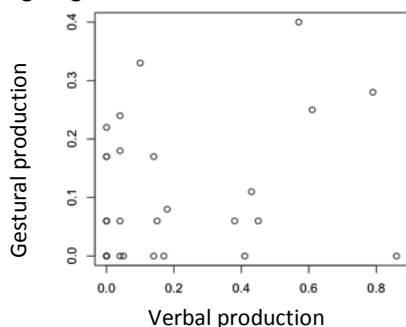
Figure 1

References. [1] Slobin. (1987). Papers from the 13th Annual Meeting of the Berkeley Linguistics Society. [2] Ferreira & Slevc. (2007). The Oxford handbook of psycholinguistics, 453–469. [3] Tanaka et al. (2011). JML, 65, 318–330. [4] Goldin-Meadow et al. (2008). Proceedings of the National Academy of Sciences, 105(27), 9163-9168. [5] Langus & Nespor (2010). Cognitive Psychology,60, 291-318. [6] Matzar et al. (1999). Gramatica del Idioma Kaqchikel.

Comparing measures of word confusability and their effect on speech production

Esteban Buz & T. Florian Jaeger (University of Rochester)
ebuz@bcs.rochester.edu

To what extent are the systems underlying language production organized to facilitate efficient information transfer? According to one position, words are articulated with more signal when they are *confusable* with other words [1, cf. 2]. According to others, the phonetic realization of words is predominantly or exclusively determined by *ease of articulation* [3-5]. In comparing these hypotheses, researchers have relied on a variety of measures of confusability. Previously used measures make the assumption that all phonemes are equally confusable. We first compare two of these measures against each other and a novel measure of word confusability based on a database of perceptual word-to-word confusability. We then compare the predicted effect of these measures on word durations from a picture-naming database.

Log word durations were derived from a picture naming study (40 target words; 60 fillers; 36 participants). Each target had one neighbor occur in the experimental list to add to potential confusability. Targets had a large continuum of confusability. We compare **three operationalizations of confusability**: (1) lexical neighbor count (*NHD*); (2) the sum of all lexical neighbors' frequencies (*fNHD*); (3) the frequency-weighted probability of confusing a word with any neighbors (*fCON*) [cf. 6] based on word-to-word perceptual confusability matrices [7]. In all measures a neighbor different from the target word by one phoneme. A version of each measure was derived from each of three **lexical databases**: MRC2 [8], CELEX2 [9], and SUBLEXus [10]. Add-1-smoothed frequency counts were used for the two frequency-weighted measures. All measures were log-transformed.

Comparisons of the same measure across databases yielded high correlations (.80<*r*<.90). **Comparisons of different measures within the same database** yielded high correlations for NHD and fNHD (.86<*r*<.92), but low correlations of those measures with fCON (.31<*r*<.47). More confusable words were associated with *shorter* duration regardless of **measure and database**. This effect reached significance (1) for NHD in all databases, (2) for fNHD in MRC2 and CELEX2, and (3) for fCON in SUBLEXus. All effects held when frequency was controlled for. This effect suggests that these measures of confusability do not affect *speech rate*, at least not in monologues (non-interactive picture description). Instead, speech rate was facilitated for words with multiple neighbors. Remarkably, binning the targets (high vs. low confusability) yielded the opposite result, suggesting an explanation for conflicting results in the literature ([11, 12] vs. [5]).

References
1. Lindblom, B., in *Speech Production and Speech Modeling*, W.J. Hardcastle and A. Marchal, Editors. 1990, p. 403-439.
2. Galati, A. and S.E. Brennan. Journal of Memory and Language, 2010. **62**: p. 35-51.
3. Bard, E.G., et al. Journal of Memory and Language, 2000. **42**: p. 1-22.
4. Arnold, J.E. Language and Cognitive Processes, 2008. **23**: p. 495-527.
5. Gahl, S., Y. Yao, and K. Johnson. Journal of Memory and Language, 2012. **66**: p. 789-806.
6. Luce, P.A. and D.B. Pisoni. Ear and hearing, 1998. **19**: p. 1-36.
7. Woods, D.L., et al. The Journal of the Acoustical Society of America, 2010. **127**: p. 1609-23.
8. Wilson, M. Behavior Research Methods, Instruments, & Computers, 1988. **20**(1): p. 6-10.
9. Baayen, R.H., R. Piepenbrock, and L. Gulikers, *CELEX2*, 1996, Linguistic Data Consortium: Philadelphia.
10. Brysbaert, M. and B. New. Behavior research methods, 2009. **41**: p. 977-90.
11. Munson, B. and N.P. Solomon. Journal of speech, language, and hearing research, 2004. **47**: p. 1048-58.
12. Scarborough, R., in *Laboratory Phonology 10*, C. Fougeron, et al., Editors. 2010, p. 557-5

Poster Session 1

Structure selection during sentence production: A role for executive control?
Maartje van de Velde, Agnieszka E. Konopka, Antje S. Meyer (MPI for Psycholinguistics, Radboud University)
maartje.vandevelde@mpi.nl

Multiple syntactic alternatives are often available to express one message. One of the factors driving the choice for a syntactic frame is verb bias. This study focuses on the role of verb bias in the process of selecting a syntactic frame for dative sentences. While some verbs are typically used with one structure (e.g., *voorleggen [submit]* and the prepositional object dative in Dutch), other verbs have a weaker bias towards one syntactic frame (e.g., *voorstellen [propose]*): the latter can be used interchangeably in the prepositional object dative (PD) and double-object dative (DO) construction, and thus allows for some degree of syntactic flexibility during production. On one view, syntactic flexibility may facilitate production because it enables speakers to fill the post-verbal sentence slots with either a direct object or an indirect object (the *incremental* view), while on a different view, flexibility can lead to competition between structural alternatives, delaying the production of the sentence until this competition is resolved (the *competition* view)[1]. The two views make opposite predictions regarding the production of sentences featuring verbs with different biases. The *incremental* view predicts shorter verb onsets for sentences featuring weak-bias verbs than strong-bias verbs, while the *competition* view predicts shorter onsets for sentences with strong-bias verbs. In addition, if the competition view holds, sentence production may benefit from a mechanism that helps resolve competition between two syntactic frames by suppressing one frame to enable fast selection of the other frame. We hypothesized that executive control (EC) can mediate this selection process, facilitating structure selection in the weak verb bias condition.

The current study compared the speed of producing Dutch datives featuring verbs with a strong and a weak bias towards the double-object and prepositional dative structure. 28 strong-weak bias verb pairs were selected from a corpus analysis of the Dutch dative alternation[2] (controlling for verb frequency, length and number of syllables). For each verb pair, one sentence was constructed which could accept both verbs (e.g., strong bias: *The student submits the plan to the professor* vs. weak bias: *The student proposes the plan to the professor*) in both a DO and PD sentence frame. Target sentences were intermixed in a list of intransitive filler sentences. All sentences were presented to participants (*n*=36) in an RSVP paradigm (Rapid Serial Visual Presentation): on each trial, speakers read a sentence presented rapidly one word at a time, performed a short distractor task, and then saw a sentence preamble (e.g., *The student…*) which they had to complete with the sentence they had just read. This paradigm allows elicitation of sentences with the target verbs and nouns, while still simulating a production situation: fast presentation of the sentences promotes reconstruction, rather than verbatim repetition of the sentence after the preamble. Participants also completed the Flanker task[3], an EC task commonly used to assess the ability to selectively inhibit one response (here: the alternative frame) in order to select a competing response (here: the target frame).

Onset analyses were performed on sentences reconstructed with the original wording and structure (66% of all responses, or 665 sentences). In line with the *competition* view, participants were 37 ms faster to initiate sentences featuring strong-bias verbs than weak-bias verbs, $β = -35.28$, $t = -2.03$. In addition, participants with higher EC were only 10 ms faster on strong-bias than weak-bias sentences, while participants with poorer EC showed a larger verb bias effect and were 63 ms faster to initiate strong-bias sentences in both frames, $β = -229.05$, $t = -1.66$, $p = .10$. This suggests that EC can facilitate structure selection in sentences with high syntactic flexibility. In contrast to earlier studies of syntactic flexibility with dative verbs[1], the results provide a first indication that equipotent syntactic alternatives can compete during sentence production and suggest that EC can play a role in resolving this competition.

References
1 Ferreira (1996). *JML, 35,* 724–755.
2 Colleman (2009). *Lang Sci, 31,* 593-611.
3 Eriksen, & Eriksen (1974). *Percept Psychophys, 16*, 143–149.

Effects of animacy on processing relative clauses in older and younger adults

Gayle DeDe (University of Arizona)
gdede@arizona.edu

Rayner, Reichle, Stroud, William, and Pollatsek (2006) suggested that normally aging adults adopted riskier processing strategies than younger adults when reading sentences with high and low frequency words. They claimed that older adults tried to "guess" upcoming words in a sentence, and experienced additional processing costs when they were wrong. This type of strategy can be viewed as "smart" in that it involves reliance on experience-based expectations to compensate for age-related changes in sentence processing ability. The present experiment tested the hypothesis that older adults also use experience-based expectations to facilitate comprehension of syntactically complex sentences (e.g., object relatives).

There is evidence that the preference for animate subjects in English influences how college age adults process sentences with object relative clauses. Traxler, Morris, and Seely (2002) demonstrated that the processing cost associated with object relatives is influenced by the animacy of the sentential subject. They claimed that readers initially interpreted the sentential subject in sentences like (a - d) as the agent of the relative clause, but were able to abandon this interpretation more easily when the subject was inanimate than animate. As a result, reinterpreting the sentential subject as the object of the embedded verb in sentences like (b) and (d) was easier when the sentential subject was an inanimate, rather than animate, noun. If older adults rely on knowledge-based expectations regarding "good" and "bad" agents to a greater extent than younger adults, then they may experience greater processing disruptions when those expectations are violated.

    (a) The musician that **witnessed** the accident **angered** the policeman a lot. (SR-AI)
    (b) The musician that the accident **terrified angered** the policeman a lot. (OR-AI)
    (c) The accident that **terrified** the musician **angered** the policeman a lot. (SR-IA)
    (d) The accident that the **musician** witnessed **angered** the policeman a lot. (OR-IA)

Older and younger adults (n=24 per group) listened to sentences with subject and object relative clauses in a self-paced listening task. Older adults were 60 to 85 years of age, and younger adults were 18 to 30 years of age. The materials (see examples a-d) were taken from experiment 3 in Traxler et al., 2002. The sentences contained animate or inanimate subjects, with inanimate or animate nouns inside the relative clause. The relative clause and main verbs are bolded in examples a-d.

The interaction between group, animacy, and sentence type was significant for both verbs (F's $\geq$ 3.88, p's $\leq$ .05). Both older and younger adults had longer listening times (LTs) for object than subject relative clause verbs in both animacy conditions (OR-AI vs. SR-AI & OR-IA vs. SR-IA). Listening times for the relative clause verb in the OR-AI condition were longer for older than younger adults, but the age groups did not differ in the other conditions. At the main verb, older adults' LTs were longer for OR-AI sentences than SR-AI or OR-IA sentences. There were no other significant differences for either age group.
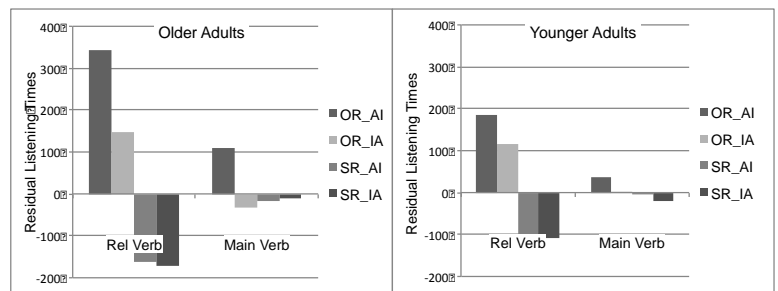
The results generally replicate those reported by Traxler et al. (2002) in the auditory modality and using a different task. More importantly, the results suggest that the animacy manipulation disrupted older adults more than younger adults. This is consistent with the claim that older adults rely on experience-based expectations to a greater extent than younger adults. Although Rayner et al. (2006) called this strategy risky, it is likely an effective compensatory strategy for age related declines in sentence comprehension in many linguistic situations.

Effects of syntactic complexity in an incremental sentence/sentence dual task
Joshua Levy (University of Massachusetts, Amherst) & William Evans (Boston University)
jwlevy@psych.umass.edu

Introduction. One long-standing issue in the literature is whether sentence processing utilizes a general verbal working memory system (e.g., King & Just, 1991) or instead relies on a specialized set of resources (e.g., Caplan & Waters, 1999). Fedorenko et al. (2007) addressed this question using an incremental dual task approach in which subjects performed incremental reading and arithmetic tasks simultaneously. The authors reported super-additive increases in RT in contexts where difficult syntax and arithmetic tasks were encountered simultaneously, and therefore argued via Additive Factors Logic (Sternberg, 1969) that reading and arithmetical operations relied on shared working memory resources. However, recent work by Evans, Caplan, and Waters (2011) failed to replicate these RT effects in two experiments, which casts potential doubt on either the strength of the relationship or the validity of the incremental dual task methodology itself. The current study was designed to test the latter using a similar incremental dual task design, but controlled for potential resource differences by presenting two tasks of the same type; in this case, two sentences to be read. If the original logic holds, then the simultaneous presentation of two syntactically complex sentences should exceed available cognitive resources and result in a super-additive increase in processing difficulty compared to the presentation of two simpler sentences or a complex sentence paired with a simple sentence.

Study design. This claim was tested by conducting two separate experiments. Expt 1 manipulated syntactic complexity by contrasting reversible subject- and object-extracted relative clause sentences ('easy' SRC vs. 'hard' ORC), while expt 2 contrasted reversible actives and reduced relatives ('easy' A vs. 'hard' RR). Items from both experiments were presented in the same list. On each trial, subjects were presented with two sentences in parallel (with one sentence directly above the other) using a phrase-by-phrase moving window self-paced design. At each key press, the next window of each sentence appeared, and at the end of each trial, one verification question was asked either about the top or bottom sentence. Trials were presented in one of three conditions: a) two 'hard' sentences, b) two 'easy' sentences, or c) one sentence of each. In all sentences, the second window was the critical region containing the manipulation of syntactic complexity. To minimize lexical interference in memory as much as possible, one sentence always contained animate nouns while the other contained inanimate nouns. NP order and presentation order were counterbalanced, as were the animacy, complexity, and question targets of the top and bottom sentences. Subjects saw a total of 12 trials for each condition, for a total of 36 trials in each experiment.

Expt 1 item	The trumpeter	who loved the drummer	formed the band	two years ago.
(SRC/ORC)	The carriage	that the truck pulled	passed the school	at night.
Expt 2 item	The lies	exposed by the forgery	harmed the defendant	at the trial.
(RR/A)	The tenant	scolded the maid	who left the apartment	in poor condition.

Results. A total of 48 subjects were tested to match sample sizes from Fedorenko et al. (2007) and Evans et al. (2011). Analyses were run using linear and logistic mixed-effects models with crossed random subject/item slopes and intercepts. In both experiments, analysis of reading times in the critical window revealed main effects of syntactic complexity in both the top and bottom sentence positions, with longer RTs associated with ORC compared to SRC conditions, and longer RTs associated with RR compared to A. Critically, there was no super-additive increase in reading time associated with the presentation of two syntactically complex sentences. Also, there were no significant effects of condition on question accuracy in either experiment.

Conclusion. These data do not support the predicted super-additive interaction in reading times in contexts where two syntactically complex sentences are read simultaneously. Since it may be assumed that both sentences draw on the same processing resources, this casts doubt on the sensitivity of incremental dual task designs of this type when investigating the nature of cognitive resources engaged in typical sentence processing.

Verbal WM capacities in sentence comprehension: Evidence from aphasia

Yingying Tan, Randi Martin (Rice University) & Julie Van Dyke (Haskins Laboratories)
yt3@rice.edu

Successful sentence comprehension often requires the ability to link non-adjacent constituents. Prior studies with healthy individuals implicate an associative, direct-access mechanism in this operation (McElree et al, 2003). Such findings have been taken as evidence against a role for working memory capacity in sentence comprehension with individual differences instead deriving from variation in linguistic abilities (e.g., vocabulary; Van Dyke & Johns, 2012). However, prior studies with aphasic patients support a role for semantic (but not phonological) STM in maintaining word meanings prior to integration (Martin et al., 1994). A recent study with healthy young subjects found evidence for a role for semantic but not phonological STM in resolving interference from semantically related nouns (Tan et al., 2011). Neither STM capacity was related to the ability to resolve syntactic interference, though reading span was related. The results were interpreted as suggesting separable capacities for maintaining semantic and syntactic information. The present study examines interference resolution in aphasic patients, who show dramatic variation in their semantic and phonological STM capacities.

Semantic and syntactic interferences were manipulated in a 2 x 2 design by varying the semantic plausibility of the intervening noun as the subject of the main verb and varying the syntactic role of the intervening noun (subj vs. obj; cf. Van Dyke, 2007). Both manipulations have been shown to increase difficulty in linking the head noun to the main verb.

Example (Semantic manipulation in brackets):
LoSyn: The jockey who had challenged the unbeatable [record/champion] yesterday will win.
HiSyn: The jockey who claimed that the [record/champion] was unbeatable yesterday will win.

Eight aphasic patients with good single word reading comprehension abilities but different patterns of performance on STM tasks and seven controls were assessed on whole sentence reading times and on time and accuracy to answer comprehension questions (e.g. Will the jockey win?"). Interference effects were calculated by computing difference scores for RT and accuracy for low vs. high interference conditions. For log reading times, the semantic and syntactic interference effects for each patient were within the range of controls. However, most of the patients showed exaggerated sensitivity to either semantic or syntactic interference for comprehension questions in both RT and accuracy. Results of linear mixed-effects models revealed a significant interaction between category probe span and the size of the semantic interference effect for accuracy, $p = .03$. However, weights for the interaction between phonological STM and semantic interference were not significance nor were there any interactions between syntactic interference effects and either semantic or phonological capacity. We are now collecting data from another 2 patients.

The current results are consistent with the view that semantic STM is related to ability to resolve semantic but not syntactic interference, and phonological STM is unrelated to ability to resolve either type of interference. These results could be explained on the grounds that individuals vary in the decay rate for different types of information outside the focus of attention. Poor maintenance of semantic information leads to difficulties in resolving semantic interference.

References

Martin, et al., (1994). Language Processing and Working Memory: Neuropsychological Evidence for Separate Phonological and Semantic Capacities. *JML, 33*, 83-111. McElree et al., (2003). Memory structures that subserve sentence comprehension. *JML, 48*, 67f. Tan, Y., Martin, R.C., & Van Dyke, J.A. (2011). Interference and working memory in sentence comprehension. CUNY Sentence Processing Conference, Palo Alto, CA., March. Van Dyke, J.A. (2007). Interference effects from grammatically unavailable constituents during sentence processing. JEP:LMC, *33(2),* 407-430. Van Dyke, J.A., & Johns, C.L. (2012). Memory interference as a determinant of language comprehension. *Language and Linguistic Compass*, 6(4), 193-211.

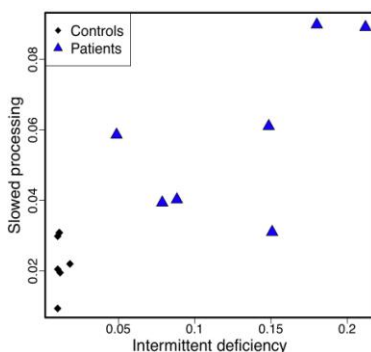**Modeling individual differences in processing deficits in aphasia**

Umesh Patil, Sandra Hanne, Shravan Vasishth & Frank Burchert (University of Potsdam)
umesh.patil@gmail.com

Caplan et al. (2007) have proposed that the sentence processing deficit in agrammatic aphasia is a consequence of *intermittent deficiency* (ID) in the capacity to carry out syntactic, semantic, and task-related computations. We operationalize ID as a specific impairment in a cognitive architecture for sentence processing. We also implement a complementary hypothesis of *slowed processing* (SP), which ascribes the deficit to a pathological slowdown (Hanne et al., 2011). We contrast these processing deficit accounts with a representational deficit account, the *Trace Deletion Hypothesis* (TDH) (Grodzinsky, 2000), which claims that patients suffer from impairment in their syntactic representation.

We implemented computational models of the two classes of accounts in a cue-based retrieval architecture of sentence processing (Lewis & Vasishth, 2005). We modeled the sentence-picture matching task reported in Hanne et al. (2011). In this study, patients (n=7) and aged-matched controls (n=7) listened to German canonical (Der Sohn fängt den Vater 'The son is catching the father') and analogous reversible non-canonical (**Den** Sohn fängt **der** Vater 'The father is catching the son') sentences while they were shown two pictures (a target and a distractor) on the screen. After the sentence ended, the participants selected the picture matching the sentence. The data consists of eye movements during the sentence presentation, picture selection responses, and response times. We first modeled controls' offline responses (accuracy and response time) and then induced impairments in the model depending on the hypotheses: ID as increased utility noise in the system (utility noise defines the degree of nondeterminism in selecting processing rules), SP as slowed procedural memory, and the TDH as an absence of trace information in the parse tree. To characterize the between subject variance in aphasia, we estimated different values for the two parameters (utility noise and slowed procedural memory) for each patient.

We evaluated the models using the root-mean-square deviation measure. The models with processing deficits (ID and SP) outperformed the models with a representational deficit (TDH). In fact, assuming only one of the two processing deficits was inadequate to capture all the crucial patterns in the data. The additive model captured:(1) the reduced accuracy in the picture matching response, (2) chance-level performance on non-canonical sentences, (3) high response times, and (4) divergent eye movements for correct and incorrect responses. Moreover, the model captured the high variability between patients' responses–different patients showed different degrees of intermittent deficiency and slowed processing, and they formed a more spread out distribution of parameter values than that for controls (see the figure).

Figure: Variability between participants across two processing impairments

In sum, the modeling results support the hypotheses of processing impairments as the source of the sentence processing deficit in agrammatic aphasia, and also demonstrate the differences at the level of individual patients in the severity of intermittent deficiency and slowed processing.

Parasitic gaps inside subject islands in (non-)native sentence processing: Evidence from eye movements during reading

Oliver Boxell & Claudia Felser (University of Potsdam)
boxell@uni-potsdam.de

Island domains are sentence regions within which fronted constituents ('fillers') cannot usually find any legitimate associated gaps. One type of exception are so-called parasitic gaps (PGs) (Engdahl, 1983). A filler can be linked to a potential gap inside certain island domains - such as the bracketed subject island in (1) below - provided that it can also be linked to a normal licit gap outside of the island domain.

   (1) What did [SUBJECT the attempt to fix __] ultimately break __ ?

Phillips (2006) showed that native English readers postulated PGs online inside subject islands containing non-finite clauses but not inside finite clauses. The present study further investigates the real-time processing of structures similar to (1) and had three main objectives. First, we sought to examine the processing time-course of PG structures and of the island domain in which they occur by using eye-movement monitoring, a method which provides more fine-grained reading-time records. Secondly, we investigated how the possibility of postulating a PG inside an island might affect the way the 'real' *wh*-dependency is ultimately resolved. Thirdly, we examined whether online sensitivity to PG environments extends to non-native readers, who have thus far only been shown to be sensitive to island domains that do not permit PGs (Felser et al., 2012).

To this end, we carried out an online reading task and an offline scalar judgment task with 24 native and 24 proficient non-native speakers of English with German as their first language. The results from the judgment task confirmed that subject island environments must be non-finite in order to host a PG, replicating earlier findings by Phillips (2006) and extending them to non-native speakers. The materials for the online reading task were modelled after Phillips (2006). Finiteness and the plausibility of the filler as a direct object of the verb inside the subject island were manipulated in a 2x2 design, as illustrated by (2) below.

   (2) The policeman knew which {prisoners/houseplants} [the activities {to inspire/that inspired} massively and cleverly the sensible criminals who want a life inside of the law] would help_ because the training programmes available included a gardening club.

At the postverbal adjunct (*massively and cleverly*) the two participant groups showed different reading time patterns in 1st pass reading times. Only the native group showed the expected Finiteness x Plausibility interaction (t=2.09), with plausibility effects restricted to non-finite environments (as in Phillips, 2006). The non-native group, in contrast, showed shorter reading times in the plausible than in the implausible direct object conditions irrespective of finiteness, and faster overall reading times for the infinitival conditions. This is indicative of an early general plausibility evaluation independent of finiteness in the latter group. A significant Finiteness x Plausibility interaction was observed in rereading times (t=2.193) that was not modulated by Group, with effects of plausibility seen only in infinitival environments. Significant finiteness by plausibility interactions were also found at the real gap (the verb *help*) in 1st pass (t=2.227) and rereading times (t=2.312).

Our findings show that (i) both native and non-native speakers are sensitive to PG environments during processing; (ii) application of the subject island constraint seems to be delayed in non-native speakers; and (iii) the attempt to link an implausible filler to a potential PG also impedes filler integration at the licit gap site.

References
Engdahl, E. (1983) Parasitic gaps. *Linguistics and Philosophy* 6, 5–34; Felser, C., I. Cunnings, C. Batterham & H. Clahsen (2012). The timing of island effects in nonnative sentence processing. *Studies in Second Language Acquisition* 34, 67-98; Phillips, C. (2006). The real-time status of island phenomena. *Language* 82, 795–823.

Poster Session 1

Aspectual interpretation and increment size: A cross-linguistic eyetracking study
Oliver Bott & Anja Gattnar (SFB 833, Tübingen University)
oliver.bott@uni-tuebingen.de

Languages differ in their grammatical means of expressing aspect (e.g. Comrie 1976) but what are the processing consequences thereof? We conducted a cross-linguistic eyetracking study to investigate how online processing is adapted to the aspectual properties of a language. We compared the time course of aspectual mismatch detection in German which lacks overt marking of aspect with mismatch detection in Russian. We hypothesized that in an aspect language like Russian temporal adverbials immediately impose restrictions on aspect choice, whereas non-aspect languages leave the aspectual interpretation underspecified until the processor has encountered a complete predication, i.e. the verb and its arguments (**C**ross-linguistic **A**spectual **V**ariation (**CAV**) Hypothesis). Thereby, the processor minimizes the risk of aspectual reanalysis in a non-aspect language like German. We exploited the relatively free word orders of German and Russian and constructed 36 items, embedded in larger sentence contexts, in a 2x2 within design: (A) aspectually mismatching adverbials (*for x time*) vs. controls (*x time ago*) appeared (B) either before or after transitive achievements (e.g. *the famous boxer$_{Subject}$ won$_{Verb\ (Russian:\ perfective)}$ the fight$_{Object}$*). We tested SVO-Adv vs. Adv-VSO order in German and SVO-Adv vs. Adv-VOS order in Russian. The **CAV** hypothesis led us to expect an immediate mismatch effect at the verb in Russian Adv-VOS sentences, whereas German readers of Adv-VSO sentences should wait until the object before being able to detect an aspectual mismatch.

Pretest. Russian sentences of the type *for two hours won$_{perfective}$* cannot be continued in any sensible way. This is different for their German counterparts as in *zwei Stunden lang gewann niemand etwas* (*for two hours won nobody anything*). Finding delayed mismatch effects in German Adv-VSO sentences as opposed to immediate effects in Russian could thus be due to anticipating an appropriate continuation. We conducted a combined acceptability judgment/continuation study. 60 native German speakers had to decide whether up to the verb the items in the Adv-V(OS) mismatch condition allowed for a sensible continuation and if so, had to write it down. In the majority of the cases (70%) participants rejected the sentence beginnings as nonsensical. Thus, immediate effects could in principle be expected in both languages.

Exp. 1. 36 monolingual native German speakers read the items (+112 fillers) and provided a sensicality judgment after each sentence. Mismatch was correctly rejected in 82% of the trials indicating that readers paid attention to aspectual properties. In the SVO-Adv order mismatch slowed down RTs of the adverbial during first-pass reading (mismatch: 702ms vs. control: 581ms; ANOVAs $p_{1/2}<.01$) and regression path durations (930ms vs. 684ms; $p_{1/2}<.01$), but did not lead to a significant increase of regressions out. As predicted by the **CAV** hypothesis, in the Adv-VSO order mismatch and control differed neither in first fixation durations, first-pass times, regression path durations or proportions of regression out before the object. Only then mismatch started to yield effects resulting in more regressions out (21% vs. 13%; GLME: p<.05) and longer regression path durations (652ms vs. 534ms; $p_{1/2}<.01$) than in the control condition.

Exp. 2. 24 monolingual native Russian speakers experienced the setup from Exp.1. In the SVO-Adv word order we observed clear mismatch effects at the adverbial: mismatch increased the proportions of regressions out (40% vs. 13% (control); GLME: p<.01) and led to increased regression path durations (mismatch: 1175ms; control: 875ms; ANOVAs $p_{1/2}<.01$). In the Adv-VOS order mismatch yielded immediate effects at the object immediately following the achievement verb with the perfective suffix: It led to more regressions out (38% vs. 23% (control); p<.05) and longer regression path durations (mismatch: 770ms; control: 587ms; $p_{1/2}<.01$) than control. The effects appeared locally at the word following the aspectually marked verb. At the following two regions containing the subject (*the famous$_1$ boxer$_2$*) mismatch and control did not differ reliably.

Conclusions. German readers specified lexical aspect only after having seen the verb and both arguments. In line with the CAV hypothesis they seem to require a sentence-sized processing domain before they can start compositional aspectual interpretation. Russian native speakers, by contrast, immediately compose the adverbial with the verb in an incremental fashion.

Dissociating reanalysis and semantic reinterpretation during garden-path recovery
Gunnar Jacob & Claudia Felser (University of Potsdam)
gunnar.jacob@uni-potsdam.de

Previous research on the processing of temporarily ambiguous sentences has shown that an initially-activated incorrect interpretation often remains activated even after processing of the sentence is complete (e.g. Christianson, Hollingworth, Halliwell, & Ferreira, 2001; van Gompel, Pickering, Pearson, & Jacob, 2006). However, it is controversial whether semantic persistence is caused by a failure of syntactic reanalysis (Christianson et al., 2001), or whether reanalysis is successful but the previously activated incorrect analysis leaves a memory trace which potentially influences end-of-sentence reconstruction processes (Kaschak & Glenberg, 2004). Sturt (2007) reports that sentence-final semantic consistence affected participants' second-pass reading times at the syntactic disambiguation region in relatively easy garden-path sentences such as *The explorers found the South Pole was actually impossible to reach*, despite apparently successful reanalysis.

We report the results from an eye-movement monitoring experiment which further explores garden-path recovery in sentences containing separate syntactic and semantic disambiguation cues. 28 English native speakers read sentences which contained temporary subject-object ambiguities, such as (1) below.

(1) While the gentleman was eating(,) the burgers (which were really tasty) were still being reheated in the microwave.

We manipulated the factors Ambiguity (ambiguous/unambiguous) and NP Length (short/long) in a 2x2 design. In the ambiguous conditions, the noun phrase *the burgers* is likely to initially be considered the direct object of the preceding verb *eating*. The auxiliary *were* provides syntactic information revealing that *the burgers* is actually the subject of the upcoming main clause, while the participle *reheated* contains semantic information inconsistent with a direct-object interpretation. The unambiguous conditions contained a comma after *eating*, ruling out the direct-object analysis.

We found significant main effects of Ambiguity at the syntactic disambiguation (*were still*) in first-pass reading time, regression-path duration, re-reading time, and total time (all p's < .001), with longer reading times for the ambiguous than for the unambiguous conditions. A similar pattern occurred at the semantic disambiguation (*reheated*) in regression-path duration, rereading time, and total time (all p's < .01), but not in first-pass reading time. For the segment in between the two (*being*) we found no such effects, thus ruling out the possibility that the effects seen at the semantic disambiguation were mere spill-over effects. The same pattern of results occurred for both the short and the long versions of the sentences.

Taken together, our results suggest that the parser initiates a syntactic reanalysis immediately when encountering the disambiguating auxiliary. However, if this reanalysis were always successful, subsequent effects of ambiguity at the semantic disambiguation should not have occurred. We conclude that the garden-path effects seen at the semantic disambiguation are best accounted for by the failed-reanalysis hypothesis, and are also consistent with dual-route models of processing which allow for syntactic and semantic representations to be computed independently.

References

Christianson, K., Hollingworth, A., Halliwell, J., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology, 42,* 368-407.
Kaschak, M. P., & Glenberg, A. M. (2004). This construction needs learned. *Journal of Experimental Psychology: General, 133*, 450-467.
Sturt, P. (2007). Semantic re-interpretation and garden path recovery. *Cognition, 105(2)*, 477-488.
Van Gompel, R.P.G., Pickering, M.J., Pearson, J., & Jacob, G. (2006). The activation of inappropriate analyses in garden-path sentences: Evidence from structural priming. *Journal of Memory and Language*, *55*, 335-362.

Filler complexity in wh-extractions from islands and non-islands

Constantin Freitag & Sophie Repp (Humboldt-Universität zu Berlin)
constantin.freitag@hu-berlin.de

In a number of experiments Hofmeister & Sag (2010) and Hofmeister (2011) found that the processing of filler-gap dependencies with an extracted wh-phrase in English is sensitive to the complexity of the *wh*-phrase. Reading times (RTs) of the word after the extracted *wh*-phrase were longer after a *which*-phrase than after a bare *wh*-word. At the gap site, in contrast, RTs were shorter for '*which*-gaps' than for 'wh-gaps'. The advantage of the more complex phrase at the retrieval site was suggested to be due to a reduction of the likelihood of misanalysis ('early gap filling') thus preventing continued reanalysis, and/or to an easier retrieval because of increased activation and resistance to interference. ♦ We present findings from German that call the generality of these findings and their interpretation into question. We conducted a self-paced reading study with word-by-word presentation. 48 participants saw 36 matrix questions with an extracted *wh*-phrase that were preceded by a context sentence, as well as 80 unrelated fillers. We varied 2 factors in a 3x2 design. Factor 1 is SENTENCE TYPE (BASE = non-island extraction, CNP = extraction from a complex NP island). Factor 2 is WH-COMPLEXITY (BARE, WHICH, ADJ): in ADJ there is a *which*-phrase with a modifying adverb plus adjective so that the *wh*-phrase in ADJ is syntactically and semantically more complex than in WHICH, where the *wh*-phrase is more complex than in BARE. After the presentation of the question, participants chose 1 of 3 fragment answers, which were either correct, similar to the correct answer, or dissimilar. ♦ Statistical analysis of the RTs revealed differences for *wh*-complexity. In the matrix clause of BASE and CNP, RTs were longer in ADJ than in BARE/WHICH up to three words after the extracted phrase, which suggests that only the syntactically & semantically most complex *wh*-phrase requires higher coding costs, and that the increased cost is spread over a longer period of time than in the earlier studies. In the retrieval region in the embedded clause (main verb & aux) RTs also were longer for ADJ than for BARE/WHICH in both CNP and BASE. Thus, unlike in the above-mentioned studies on English, in the present study retrieving a more complex *wh*-phrase is more and not less costly than retrieving a less complex *wh*-phrase. For the answers there were main effects of sentence type and of *wh*-complexity in the reaction times. The correct answer was given faster for ADJ than for WHICH/ BARE. BASE was answered faster than CNP. ♦ We suggest that the disadvantage of a very complex *wh*-phrase at the retrieval site during reading might be due to the activation of more semantic features, which, as is indicated for the advantage of ADJ in the answers, helps choosing the associated information in the answer (exclusion of alternatives). Thus, we might be dealing with a trade-off here due to the task that required the processing of fairly complex information about a referent introduced in the context that later had to be used for the choice of answer. ♦ In an acceptability judgement experiment with the same materials (context + question) without an answer task we found that CNP sentences are more acceptable with ADJ than with WHICH than with BARE. There was no effect of *wh*-complexity for BASE sentences, which overall were judged to be better than CNP. This suggests (confirming results reported by H&S 2010) that the availability of more semantic features improves the acceptability of CNP. The role of complexity for the acceptability of non-islands, however, could not be confirmed.

1		*Jim will explain that Lisa has*	Wen_{who.acc}	wird_{will} Jim darle-
2	BASE	*incriminated the tenant that*	Welchen_{which.acc} Mieter_{tenant}	gen_{explain}, dass_{that} Lisa
3		*hasn't got a job anymore and that is outwardly neglected.*	Welchen_{which.acc} äußerlich_{outwardly} verwahrlosten_{neglected} Mieter_{tenant}	t_{wh} belastet_{incriminated} hat_{has}?
4		*Jim will explain the view that Lisa*	Wen_{who.acc}	wird_{will} Jim den_{the}
5	CNP	*has incriminated the tenant that*	Welchen_{which.acc} Mieter_{tenant}	Standpunkt_{view} darle-
6		*hasn't got a job anymore and that is outwardly neglected.*	Welchen_{which.acc} äußerlich_{outwardly} verwahrlosten_{neglected} Mieter_{tenant}	gen_{explain} dass_{that} Lisa t_{wh} belastet_{incriminated} hat_{has}?

Answers: a. the unemployed tenant   **b.** the married tenant   **c.** the moody detective

References. Hofmeister, P. & I. Sag (2010) Cognitive constraints & island effects Language 86, 366-415

Agreement violations in Arabic: Qualitative ERP differences between singular and plural subjects

R. Muralikrishnan (New York University Abu Dhabi), & Ali Idrissi (United Arab Emirates University, Al Ain)
r.muralikrishnan@nyu.edu

Verbs in Arabic generally agree with the subject in person, number and gender. However, for plural subjects, the verb agrees fully in the subject-verb (SV) order, but it only agrees in person and gender in the verb-subject (VS) order. Thus, an agreement mismatch for plural subjects does not always entail ungrammaticality, whilst it does for singular subjects. In the present ERP study, we investigated whether the processing system is sensitive to this idiosyncratic behaviour of plural subjects when processing verb agreement in the SV order. If so, qualitatively different ERP patterns should be elicited at the verb for singular versus plural subjects.

Methods: The stimuli were intransitive sentences of the form Temporal-Adverb–Subject–Verb–PP, which were (a.) either in the past or future tense, and (b.) the subjects of which were respectively singular or plural. The subject nouns were human (masculine & feminine) common nouns; all nouns occurred in their singular & plural forms (in different conditions) in the experiment. The verbs in the stimuli were either acceptable, or represented an agreement violation (person or number or gender; no multiple violations), or were violations of the tense established by the adverb. Thus there were six conditions: 3 condition types x 2 subject types. Thirty sentences were presented per participant in each condition; the agreement violation condition type consisted of 10 each of person, number or gender violations. All conditions were equiprobable. Stimuli were pseudo-randomised including fillers, and presented in a rapid serial visual presentation. Participants (28 right-handed Arabic native-speakers; 9 female) performed an acceptability judgement and a probe task following each trial.

Results and Discussion: Acceptable sentences elicited a P300 with a posterior maximum (P3b effect) when the subject was singular as opposed to plural. All violations engendered a biphasic negativity–late-positivity pattern Further, an additional Left Anterior Negativity (LAN) effect ensued for agreement violations when the subject was singular as opposed to plural. There were no ERP differences between singular and plural subjects for tense violations. Agreement violations in other languages have been consistently shown to engender a LAN effect (1,2,3), but in our data, the LAN is significant only for singular subjects (see Figure 1 for a topographic map of this effect). A possible reason could be that there is only one possible agreement pattern for Arabic singular subjects irrespective of word-order; any deviation from this leads to a morphosyntactic violation, reflected in a LAN. By contrast, since Arabic verbs agree differentially to plural subjects based on word-order, a mismatch may not always be a morphosyntactic violation, hence no LAN. Further evidence for this interpretation stems from the P3b effect for singular subjects in acceptable sentences in our data. The P3b has been observed for linguistic stimuli whenever a categorisation, such as an acceptability judgement, is necessary (4,5). It obtains when a prior expectation about the stimulus is fully met. In line with this, we argue thus: singular subjects require an identical agreement pattern in all word-orders, thus giving rise to a specific expectation. When this expectation is fully met (as it does in the acceptable sentences), a P3b is evoked. Such an expectation would be less specific for plural subjects given their idiosyncratic behaviour, thus no P3b. This in conjunction with the absence of a LAN for plural subjects, we argue, shows the differential nature of processing agreement for singular and plural subjects due to properties specific to Arabic. A follow-up study to observe the effects of each agreement feature separately is under way to further qualify this argument.

Figure 1. Topographic map of ERPs at the verb: agreement violations minus acceptable stimuli, 350-500ms.

Singular   Plural

−1.0   µV   +1.0
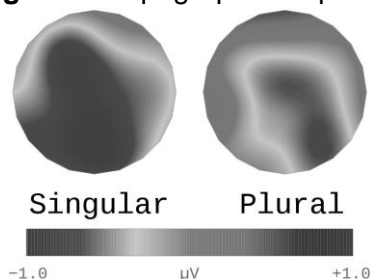
References. [1] Osterhout & Mobley (1995). *Journal of Memory and Language, 34*, 739-773. [2] Barber & Carreiras (2005). *Journal of Cognitive Neuroscience, 17(1),* 137–153. [3] Roehm et al. (2005). *NeuroReport, 16(8),* 875-878. [4] Roehm et al. (2007). *Journal of Cognitive Neuroscience 19(8)*, 1259–1274. [5] Kretzschmar (2010). Unpublished doctoral dissertation, Marburg, Germany.

Acceptability of grammatical and ungrammatical doubly nested relative clause structures in Spanish: Some evidence in favor of usage-based approaches

Florencia Reali (Universidad de los Andes)
f.reali96@uniandes.edu.co

Previous acceptability rating experiments in English showed that people find doubly nested relative clause structures just as acceptable when only two verb phrases (2 VP) are included as when the grammatically required three verb phrases (3 VP) are present (Gibson & Thomas, 1999). Subsequent studies (Christiansen & MacDonald, 2009), using materials such as (a) below, found that ungrammatical doubly embedded constructions (2 VP) were actually rated *more* acceptable than 3 VP grammatical ones. The latter findings were interpreted to be consistent with usage-based approaches since the human data were predicted by a connectionist model. Here, we provide further evidence in favor of usage-based approaches by combining corpus analysis and off-line rating tasks designed to test acceptability of doubly embedded relative clause structures in Spanish.

Pronominal object relatives are extremely frequent in English (Reali & Christiansen, 2007). Using a corpus of spoken Spanish – *Corpus lingüístico de referencia de la lengua española* – we found that, as in English, the overwhelming majority of object relatives in Spanish are pronominal (85% of 564). Moreover, Spanish allows flexibility of surface structure: embedded pronouns can be dropped (e.g., optional *nosotros* in (b) below) because subject information is encoded in verb conjugation. The corpus analysis revealed that in the vast majority of pronominal object relatives (77%) the embedded pronoun was omitted. We hypothesize that exposure to frequent "chunks" should facilitate the processing of object relatives where the embedded pronoun is omitted.

Can a subtle manipulation such as dropping the most embedded pronominal subject change grammaticality judgments in doubly nested sentences in Spanish? An off-line acceptability rating experiment was conducted to answer this question. Materials were doubly nested sentences in Spanish (adapted from Christiansen & MacDonald, 2009) that varied in two ways: they were either 3 VP or 2 VP and the most embedded subject – which was always a pronoun – was either present (low frequency condition) or omitted (high frequency condition) (see b). The results showed that participants judged ungrammatical 2 VP sentences to be more acceptable than grammatical 3 VP sentences but *only* in the low frequency condition. A 2 (NP3Present vs. NP3Omitted) X 2 (3 VP vs. 2 VP) ANOVA revealed no main effect but a significant interaction across conditions ($F1(1,24)=7.21$, $p<.02$;$F2(1,11)=7.59$, $p<.02$). Paired comparisons among grammatical and ungrammatical sentences with pronominal NP3 present (low frequency condition) showed that participants found ungrammatical 2 VP sentences more acceptable ($t1(24)=1.92$, $p=.063$; $t2(11)=2.88$, $p<.02$). Contrarily, ungrammatical 3 VP sentences *less acceptable* when the most embedded pronoun was omitted (high-frequency condition), although the difference was not significant (all p´s > 0.2). Additional paired comparisons showed that participants found 3 VP sentences with dropped embedded pronouns (high frequency condition) more acceptable than 3 VP sentences with the most embedded pronoun present (low frequency condition) ($t1(24)=2.1$, $p<.05$; $t2(11)=2.14$, $p=.055$). The significance of these findings is discussed in the light of usage-based approaches to sentence processing.

(a) The spider that the bullfrog that the turtle followed [**chased/*mercilessly**] ate the fly.
(b) La araña que el zancudo que (*nosotros*) perseguimos [**picó/*despiadadamente**] se comió la mosca.

References
Christiansen, M.H., & MacDonald M.C. (2009). A usage-based approach to recursion in sentence processing. *Lang Learn,* 59: 126-161.
Gibson, E., & Thomas, J. (1999). Memory limitations and structural forgetting: the perception of complex ungrammatical sentences as grammatical. *Lang Cognitive Proc*,*14*: 225–248.
Reali, F., & Christiansen, M.H. (2007). Word-chunk frequencies affect the processing of pronominal object-relative clauses. *Q J Exp Psychol,* 60: 161-170.

Retrieval respects (strong) crossover

Dave Kush, Colin Phillips and Jeff Lidz (University of Maryland, College Park)
kush@umd.edu

Some constraints on syntactic dependencies (e.g. Principle C, Islands) have robust and immediate effects on on-line dependency formation [1-2]. Others, (e.g. Principle B [3-5], subject-verb agreement [6-8]) are less reliably deployed on-line: many studies show interference from NPs that should be structurally irrelevant based on these constraints. The robust deployment of Principle C and islands could reflect an inherent advantage for these constraints, or it could reflect the time-course/directionality of constraint application: Principle C and Island constraints typically apply to a forwards search for an antecedent or a gap, which may avoid the use of interference-prone memory retrieval mechanisms. Here we tested these alternatives in 2 self-paced reading studies that probe the effects of Principle C in configurations that require backwards retrieval. We find that structurally inappropriate nouns do not yield interference effects, indicating that the robustness of Principle C on-line cannot be reduced to its forwards-looking nature, contra [2].

Strong Crossover (SCO) is a constraint on pronoun interpretation that has been attributed to Principle C [9]. In (1a) the filler which lady cannot corefer with the pronoun she, arguably because when the filler is interpreted as the object of the verb praise, it is c-commanded by the pronoun, creating a configuration that violates Principle C. In contrast, the filler can corefer with the anaphor herself in (1b), because the gap position is higher than the anaphor, avoiding a Crossover violation.

(1)   a. John knew which lady Bill said she praised ___ . [coreference impossible]
         b. John knew which lady Bill said __ praised herself. [coreference ok]

By comparing the online interpretation of configurations like (1a-b) we can test sensitivity to Crossover/Principle C in environments where the pronoun can engage a retrospective retrieval process. If Crossover/Principle C is respected the filler should not be considered as a potential antecedent in (1a).

Both experiments measured RTs at a pronoun within the surface scope of a wh-phrase. The pronoun (*he*) either matched (*policeman*) or mismatched (*policewoman*) the wh-phrase in gender. The wh-gap either c commanded the pronoun (NoCrossover; 2) or was c-commanded by the pronoun (Crossover; 3). If retrieval ignores Principle C/Crossover, we expect an effect of gender-match: comprehenders should be sensitive to the gender match, irrespective of Crossover. If retrieval processes respect the constraint, then we should observe a Crossover x Gender-Match interaction, with a gender-match effect only in the NoCrossover conditions.

(2) **[NoCrossover conditions]**
     Mary knew which { policeman | *policewoman} { _____ } claimed that he quickly talked to Martin after the accident.
(3) **[Crossover conditions** (Exp 2)**]**
     Mary knew which { *policeman | *policewoman} it seemed that he quickly talked to ____ …

Our results show that initial retrieval of pronoun antecedents can target a wh-filler, but only when this does not violate Principle C/SCO. This suggests that online fidelity to Principle C cannot be reduced to the use of forwards-looking search mechanisms.

References

[1] Cowart, W., & Cairns, H. S. 1987. Memory & Cognition [2] Kazanina, N., et al. 2007. JML; [3] Badecker, W. and Straub, K. 2002. JEP: LMC; [4] Kennison, S.M. 2003. JML; [5] Runner et al. 2003. Cognition; [6] Pearlmutter, N.J. et al. 1999. JML; [7] Staub, A. 2009. JML; [8] Wagers, M.W. et al. 2009. JML; [9] Chomsky, N.1981. LGB.

The interpretation of elided reflexives in children and adults
Sharese King (Stanford University), Jeffrey T. Runner (University of Rochester)
sharese@stanford.edu

Background. Our goal in this series of studies is to understand the time course of the development of the referential system in normally developing children. We focus here on the development of the pragmatic knowledge required to interpret elided "picture" noun phrases (PNPs) in an adult-like way. Previous research (Runner & Goldwater 2011) shows that adults are more likely to assign a coreferential ("strict") interpretation to an elided PNP reflexive than to an elided direct object (NP) reflexive (see (1) and (2)).

1. Dan is pointing at a *picture of himself* and Mike is too. [more coreferential (Dan) responses]
2. Dan is pointing at *himself* and Mike is too. [fewer coreferential (Dan) responses]

Experiment. We tested the development of this distribution of interpretations in two scene verification studies. In the first, we sought to replicate Runner & Goldwater by testing adults (n=26); in the second, we tested children (n=22) in two age groups: 4-6 (n=10) and 7-9 (n=12). Participants, in both experiments, listened to sentences similar to (1)-(2) while looking at a scene depicting either a bound variable (BV) interpretation for the elided reflexive (Mike pointing at himself or a picture of himself) or a coreferential (Coref) interpretation (Mike pointing at Dan or a picture of Dan). Participants responded 'yes' or 'no' to whether the displayed scene matched the sentence.

Results. The results (see figure) showed that adults and 7-9 year olds show a strikingly similar pattern of results: a significant effect of scene type (more 'yes' responses in the BV scene) and a numerical effect of NP type (more 'yes' responses for PNPs in the Coref scene). However, 4-6 year olds' responses were just above chance.

Discussion. A base-line difference between the 4-6 year olds (68%) and the older kids and adults (both 98%) in responses to sentences not containing ellipsis urges caution in interpreting the youngest kids' data. However, it does seem safe to claim that the 7-9 year olds have acquired the pragmatic knowledge to interpret elided NP and PNP reflexives in an adult-like way. The 4-6 year olds, then, may still be developing this knowledge--which seems to affect both their reflexive interpretation as well as ellipsis interpretation. This is consistent with claims in the literature that this period is when children are beginning to develop pragmatic abilities (Verbuck & Roeper 2010). Future research will focus on increasing child participation, as well as adjusting the task to make it easier for the younger kids.
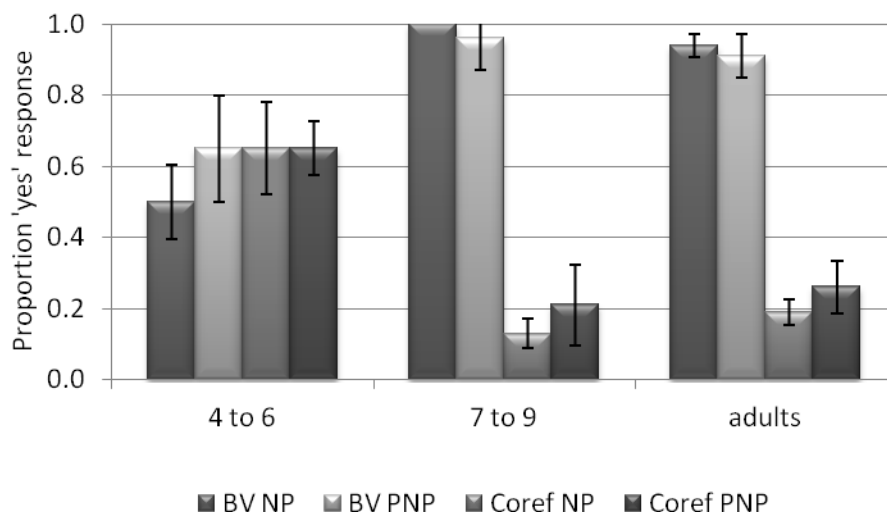
Effects of 'long-before-short' on processing of canonical and scrambled order in Japanese

Katsuo Tamaoka, Chi Yui Leung (Nagoya University) & Sachiko Kiyama (National Center for Geriatrics and Gerontology)
ktamaoka@gc4.so-net.ne.jp

Introduction. Japanese native speakers show a preference for a long-before-short noun phrase (Yamashita & Chang, 2001). If this is true, could canonical position change depending upon noun phrase length? Thus, the present study examined whether effects of phrase length can go beyond canonical order. If 'long-before-short' in Japanese is a *preference*, canonical order will play a stronger influence on sentence processing, resulting in the scrambled orders of (ii) and (iv) having longer reading times than canonical orders of (i) and (iii). If 'long-before-short' influences sentence processing, (ii) and (iv) should be processed faster than (i) and (iii). (Experiment 1)

(i)   NP-NOM (*keisatsukan-ga*) *long*NP-ACC (*kowasoona chooshin-no han'nin-o*) V (*taihoshita*)
(ii)   *long*NP-ACC$_1$ (*kowasoona chooshin-no han'nin-o*) NP-NOM (*keisatsukan-ga*) g_1 V (*taihoshita*)
(iii) *long*NP-NOM (*kowasoona chooshin-no keisatsukan-ga*) NP-ACC(*han'nin-o*) V (*taihoshita*)
(iv)   NP-ACC$_1$ (*han'nin-o*) *long*NP-NOM (*kowasoona chooshin-no keisatsukan-ga*) g_1 V (*taihoshita*)

Similar predictions were applied to active sentences of long NP-ACC or NP-DAT with ditransitive verbs under animate/inanimate conditions (Experiments 2 and 3). Using an eye-tracker (EyeLink1000), Experiment 4 examined ditransitive active sentences with both long NP-ACC/NP-DAT and canonical and scrambled orders, constructed with only animate NPs. Experiments 1-4 were conducted on native Japanese speakers

Results. (**Experiment 1** – Correct-or-incorrect decisions for active sentences with transitive verbs): (i) and (iii) were processed faster than (ii) and (iv). Long NPs had no influence on processing active sentences with transitive verbs whereas scrambling effects were apparent regardless of phrase length. (**Experiment 2** – The same task on active sentences with ditransitive verbs, long NP-ACC and other NPs are all animate): Under the condition that all NPs are animate and equal in length, canonical order was processed 261 ms faster than scrambled order. Likewise, under the condition that animate NP- DAT was longer in length, canonical order was processed 231 ms faster than scrambled order. Scrambling effect was apparent regardless of NP length. (**Experiment 3** – The same task for active sentences with ditransitive verbs under the condition of long inanimate NP- ACC): Under the condition that animate NPs and an inanimate NP-ACC are equal in length, canonical order was processed 123 ms faster than scrambled order. Likewise, under the condition that inanimate NP-ACC was longer in length, canonical order was processed as fast as scrambled order. Therefore inanimate long NP-ACC reduced scrambling effect. Yet, effects of either animacy and/or phrase length were unclear. (**Experiment 4** – Eye-tracking): Since scrambling effects were apparent when NPs were all animate, correct-or-incorrect decisions for active sentences with ditransitive verbs under the condition of animate long/short NP-ACC/NP- DAT with canonical and scrambled order was conducted using an eye-tracker. Scrambling was the major factor affecting sentence processing, while NP-length only affected when scrambled

Conclusion. 'Long-before-short' had no effect on processing active sentences with transitive verbs, but scrambling effect did. Both animacy contrast and phrase length of NP- DAT and NP-ACC reduced scrambling effects on two object NPs. When NPs are all animate, noun phrase length affects processing of scrambled sentences. Thus, scrambling effects are consistent while 'long-before-short' effects are conditional.

References. Yamashita, H., & Chang, F. (2001). "Long before short" preference in the production of a head-final language. *Cognition, 81,* B45–B55.

Resumption rescues islands after all: An experimental investigation on Italian and English

Andrea Beltrama & Ming Xiang (University of Chicago)
andremormora@uchicago.edu

Overview. While resumptive pronouns (RPs) are not grammatical in English and other languages, it has been claimed (Ross 1967, Kroch 1981) that they can be used as a last resort strategy to obviate island effects. However, recent experimental studies (Heestand et al. 2011, Han et al. 2012, Alexopoulou and Keller 2007) showed that in a variety of islands RPs are not actually more acceptable than gaps. In the current study, we assess the rescuing effect of RPs in Italian and English, and show that under certain circumstances, RPs do rescue islands.

Expt 1 Italian. Each item consisted of two different sentences. The first one provided contextual information, such as "*Yesterday there were riots in the street, and some people were wounded by the police*", while the second one contained a relative clause and came in 16 different conditions, based on a 2x2x2x2 factorial design: a) Island (Island vs Non island); b) Resumption (Gap vs Resumptive Pronoun); c) Embedding (2 vs 3 embeddings); d) Case (Accusative vs Dative on the extracted element). An island, 3-embedding, accusative case example would be

(1) This is the guy that they say that the cop who shoved him/__ had just finished his training".

64 sets of stimuli were distributed into 16 lists with a Latin Square design (40 fillers). Participants were presented the items auditorily, and after each item, they assessed how easy it was to comprehend the target sentence on a 1 to 7 scale (1= completely incomprehensible; 7 = perfectly comprehensible).

Expt 1 Results (n=43). We found a main effect of Island (p <.0001), with all island conditions being worse than non-island ones (all ps < .01). We also found a main effect of Resumption (p<.0001) and significant interaction effects between Resumption and Island (p<.001) and Resumption and Case (p<.001). Within island conditions, with accusative (but not with dative) case, RPs showed a rescuing effect, resulting in higher acceptability than gaps, both with 2-level (3.65 vs 3.24, p <.05) and with 3-level embedding (3.92 vs 3.40, p <.001). Our results suggest that Italian RPs can rescue island violations, at least in the accusative case. At the same time, the fact that resumption is never better than gaps outside of islands confirms that resumption in Italian is not fully grammatical. The question remains as to why the rescuing effect of RPs was undetectable in English in previous studies. We suggest two possibilities. First, since Italian allows RPs in in constructions such as Clitic Left Dislocation ("CLLD", Cinque 1990), RPs could be more accessible to Italian speakers than English speakers in rescuing situations too, as RPs independently exist in the Italian. Second, an additional difference between the current study and previous ones is that we provided a background context sentence for each target sentence, which could have facilitated the overall comprehension.

Expt 2 English. In this study we look at whether it is true RPs in English is unacceptable, and whether that affects how RPs can rescue islands in English. In this study, we tested RPs in both relative clause islands and topicalization (the latter is similar to Italian CLLD case). The design is largely similar to Expt1 (but with only the accusative forms), manipulating Resumption (gap vs. RP), Embedding (2 vs. 3), Island (island vs. non-island) and construction type (relative clause vs. topicalization). Stimuli were presented in written form on Amazon MTurk (n=64). Within islands, RPs were significantly better than gaps both in relative clauses and topicalization, showing a rescuing effect for English as well (Relative Clause: 4.48 vs 3.98 2-Emb and 4.29 vs 4.00 3-Emb; CLLD: 4.20 vs 3.90 2-Emb, 4.28 vs 4.00 3-Emb; all ps <.01). These results suggest that the island-independent presence of RPs in CLLD in Italian (and not in English) cannot be the reason why rescuing effects were not found in previous studies on English. We therefore propose that the discrepancy between our findings and those from the previous studies is due to the presence of a background sentence in the presentation of the stimuli. We are currently conducting a third experiment to test the effect of context in English.

A connectionist model of Mandarin relative clause processing asymmetries
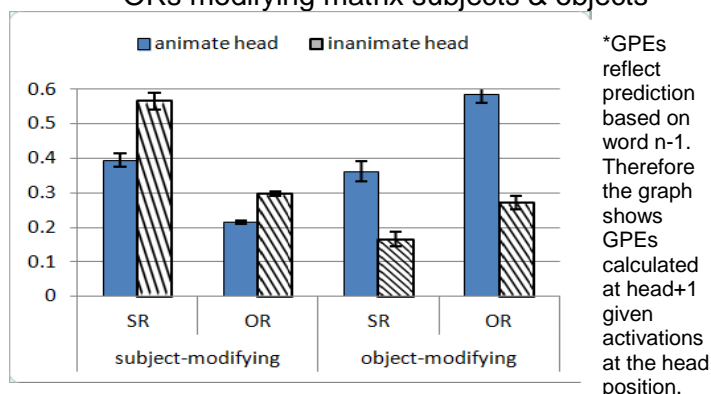Yaling Hsiao & Maryellen MacDonald (University of Wisconsin-Madison)
yhsiao5@wisc.edu

Unlike results in many other languages, Mandarin subject relative clauses (SRs, e.g. the girl who kissed the boy) may be harder than object relatives (ORs, the girl who the boy kissed). Some trace these results to Mandarin relative clauses (RCs) being head-final, with different configurations of noun-verb dependencies than in other languages (Hsiao & Gibson, 2003). Others find opposite results (SRs easier than ORs) and posit a universal SR preference independent of dependency length (Lin, 2006). Still others suggest that animacy (Wu, Kaiser & Andersen, 2012) or the matrix position of the RC head (Lin, 2006) explain SR/OR asymmetries. This diversity of materials, results, and interpretation serve to marginalize these potentially informative phenomena. We address Mandarin RC processing in a simple recurrent network (SRN), permitting us to cross more factors than is possible in a human study.

The SRN contained 29 localist input and output units and 40 hidden units with the learning rate of 0.05 and batch size of 1. It was trained with one epoch of 10K simple and RC sentences, with structural frequencies and noun animacy patterns based on the Chinese Treebank (Xue et al., 2010) (Table 1). Performance was assessed with Grammatical Prediction Error (GPE) (MacDonald & Christiansen, 2002), with a scale of 0 (accurate) to 1 (wrong). Following training, the model was tested on novel sentences crossing RC type (SR/OR), modification site (matrix subject vs. object), head noun animacy, and embedded noun animacy (as units are localist, "animacy" is distributional). A subset of the data is described here; GPEs only at the head noun are presented because a) this position follows the RC and is a main site of difficulty in reading studies and b) this is the first position for observing head animacy effects. Overall, for RCs modifying the matrix subject, SRs had higher GPEs than ORs, replicating Hsiao & Gibson (2003) (see Fig. 1), but there's also an effect of head noun animacy, replicating Wu et al. (2012). Animate heads yield lower error because they are more common as matrix subjects. These results suggest that some reading time effects that have been attributed to RC complexity may instead reflect the statistics of animate vs. inanimate nouns in matrix clauses. RCs modifying matrix objects yield a different pattern, generally consistent with Lin (2006), but with larger head animacy effects. Here inanimate heads are easier (inanimate matrix objects are more common) but there is also an interaction with RC type, reflecting the predominance of ORs with inanimate heads. We trace these complex effects to the statistics of the input (at both matrix and RC levels) and to varying degrees of ambiguity unfolding over time in the various RC-maxtrix parings. The model captured many human findings with no structural knowledge built in, suggesting the substantial effects of learning over the patterns of language use in these sentences.

Table 1. Finite state context free grammar (with corpus-based transitional probabilities)

S → subNP + VP (0.7) / VP (0.3)
VP → VI (0.21) / VT + objNP (0.79) (VI=intrans VT=transitive)
subNP → aniN(0.5) / inaN (0.34) / subjRC (0.16)
objNP → aniN(0.18) / inaN (0.73) / objRC (0.09)

subRC (modifying matrix subject)
 → SRC_VI (0.11): VI + DE + aniN(0.2)/inaN(0.8)
 → SRC_VT (0.58): VT + aniN + DE + aniN(0.47)/inaN(0.53)
       : VT + inaN + DE + aniN(0.87)/inaN(0.13)
 → ORC (0.31): aniN + VP + DE + aniN(0.14)/inaN(0.86)
       : inaN + VP + DE + aniN(0)/inaN(1)

objRC (modifying matrix object)
 → SRC_VI (0.26): VI + DE + aniN(0.31)/inaN(0.69)
 → SRC_VT (0.47): VT + aniN + DE + aniN(0.53)/inaN(0.47)
       : VT + inaN + DE + aniN(0.45)/inaN(0.55)
 → ORC (0.27): aniN + VP + DE + aniN(0.16)/inaN(0.84)
       : inaN + VP + DE + aniN(0)/inaN(1)

Figure 1. GPEs at the head position* for SRs & ORs modifying matrix subjects & objects

*GPEs reflect prediction based on word n-1. Therefore the graph shows GPEs calculated at head+1 given activations at the head position.

Poster Session 1

Effects of verb meaning on lexical integration in agrammatic aphasia

Jennifer E. Mack, Woohyuk Ji, & Cynthia K. Thompson (Northwestern University)
jennifer-mack-0@northwestern.edu

Introduction. Healthy young adults use the selectional restrictions of verbs to predict and facilitate integration of subsequent arguments (Altmann & Kamide, 1999, 2007; Boland, 2005). Previous research has shown that agrammatic aphasia involves deficits in access to selectional restrictions (Dickey & Warren, 2012; Myers & Blumstein, 2005) and in lexical integration (Swaab, Brown & Hagoort, 1998; Thompson & Choy, 2009). The present study used visual-world eyetracking to test the effects of verb meaning on prediction and integration of subsequent arguments in listeners with agrammatic aphasia and healthy older adults.

Method. Nine adults with agrammatic aphasia and ten age-matched controls viewed four-object arrays (e.g., jar, plate, stick, pencil) while listening to sentences containing either a *restrictive* verb semantically compatible only with the target object or an *unrestrictive* verb compatible with all four objects. Experiment 1 tested prediction and integration of overt arguments (e.g., *Susan will open/break the jar*), whereas Experiment 2 tested prediction of non-overt arguments (e.g., *Susan will open/break the …*). Eye movement data were analyzed using mixed-effects empirical logit regression (Barr, 2008).The critical time window for both experiments, selected using the grand mean fixation data, was the first 500 ms following verb offset.

Results. Figure 1 illustrates the proportion of fixations to the target picture in the critical windows. In Experiment 1, significant main effects were observed for verb type (restrictive>unrestrictive) and group (control>aphasic), with no interaction; the rate of increase in fixation proportion was affected by verb type (restrictive>unrestrictive) but not by group. Experiment 2 elicited a main effect of verb type, but not of group, and a significant group x condition interaction; the rate of increase in fixation proportion was affected by verb type (restrictive>unrestrictive) but not by group. Follow-up tests indicated that controls but not aphasic listeners made more target fixations in the restrictive condition.

Discussion. Unimpaired older listeners exhibited rapid use of verb information to predict (Experiment 2) and facilitate integration (Experiment 1) of subsequent arguments. In contrast with previous studies with young participants, these effects emerged after verb offset, suggesting that linguistic prediction processes may be slowed in older listeners (Federmeier & Kutas, 2005). For the aphasic listeners, the results revealed timely use verb information to facilitate integration of overt arguments (Experiment 1) but impaired prediction of non-overt arguments (Experiment 2). These findings suggest that access to selectional restrictions is vulnerable in agrammatic aphasia (cf. Myers & Blumstein, 2005), resulting in deficits in predicting upcoming arguments.

Figure 1. Proportion of target fixations in first 500 ms following verb offset (50 ms bins)
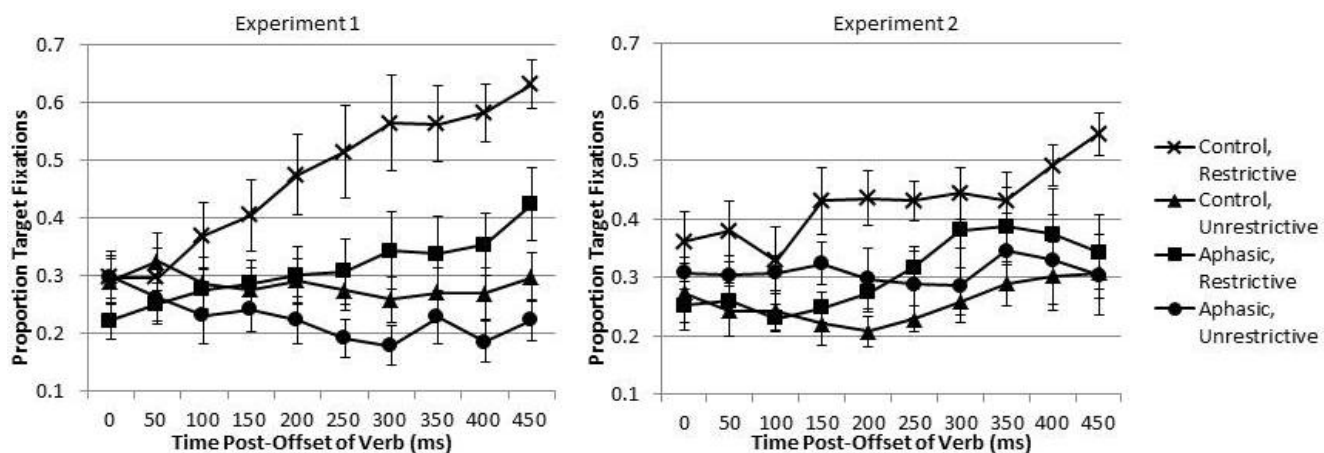
Similarity-based interference is required for the LIFG effect of object extraction: Evidence from MEG

Kimberly A. Leiken & Liina Pylkkänen (New York University)
Kal412@nyu.edu

Introduction. One of the most contentious questions in neurolinguistic literature has been the contribution of the left inferior frontal gyrus (LIFG) ("Broca's area") to syntactic processing. Perhaps the most replicated finding in this literature has been the increase of hemodynamic activity for object relative clauses (OR) (e.g. *The fireman who the deputy called saved the sailor*) compared to subject relative clauses (SR) (e.g. *The fireman who called the deputy saved the sailor*) (e.g. Caplan et al., 2000, 2008; Constable et al., 2004; Just et al., 1996; Keller et al., 2001; Stromswold et al., 1996), mirroring the behavioral finding that ORs are more costly than SRs. However, the majority of studies showing this LIFG effect have used structures with syntactically parallel noun phrases, known to invoke similarity-based interference (Lewis, 1996). In fact, the behavioral OR > SR effect has been shown to substantially diminish when the noun phrases are not parallel in surface syntax (Lewis, 1996, Gordon, Hendrick, & Johnson, 2001; Gordon, Hendrick, & Levine, 2002; Lee et al., 2007, Van Dyke & McElree, 2006). If the LIFG effect is similarly dependent on syntactic parallelism, this would suggest it does not reflect the presence of a dependency in ORs but rather similarity based interference, consistent with working memory accounts (e.g., Caplan, Alpert, & Waters, 1999; Rogalsky, Marchin, & Hickok, 2008) and cognitive control based accounts of this effect (e.g., Novick, Trueswell, Thompson-Schill, 2005). We tested this with magnetoencephalography (MEG), which also allowed us to characterize the time-course of LIFG activation.

Design. We kept our stimuli as minimal as possible and compared ORs with and without syntactic parallelism to Subject+Verb structures with no dependency. The millisecond time resolution of MEG allowed us to directly focus our analysis on the target verb, i.e., the site of the OR gap. To avoid a systematic morphological confound, half of our verbs were reflexives and thus morphologically more basic in their transitive form (e.g., Mchombo, 1993), and half causatives, morphologically more basic in their intransitive form (e.g., Pylkkänen, 2008), resulting in a 3 x 2 design with Frame and Verb Type as factors (1-3). To assure comprehension and attention, the phrases were followed by a matching task using photographs that either matched or mismatched the content of the linguistic stimulus. The stimuli were padded with consonant strings to control for the amount of visual material encountered prior to the critical verb.

1.   Subj+Verb (refl/caus):   plh rcm ycg rcm Sally bathed/walked
2.   ObjRelNonPar (refl/caus):   plh rcm the dog Sally bathed/walked
3.   ObjRelPar (refl/caus):   plh the dog the woman bathed/walked

Hypotheses. The LIFG was predicted to show an increase for both OR conditions if this effect reflects dependency formation alone. In contrast, under a similarity-based interference account, only the OR condition with syntactic parallelism should show a LIFG increase.

Results. A nonparametric cluster-based permutation test (Maris & Oostenveld, 2007) on LIFG activity identified a significant main effect of Frame in a late time window at 538-639ms (p=.0315). Planned pairwise t-tests over the significant 538-639ms cluster were performed showing a reliable increase for ObjRelPar over ObjRelNonPar (p=.0462) and for ObjRelPar over Subj+Verb (p=.0281 at 572-640ms), but no difference between ObjRelNonPar and Subj+Verb. No effects of Verb Type were identified. These results were replicated in a whole brain analysis.

Conclusion. The present findings replicate the well-studied hemodynamic LIFG effect of object extraction with a more time-sensitive technique. The findings suggest that this effect is indeed dependent on the syntactic parallelism between the two noun phrase arguments encountered before the gap site. This outcome is inconsistent with purely structural accounts of this LIFG increase but conforms well with the behavioral literature on similarity-based interference in ORs, along with the cognitive neuroscience literature on LIFG contributions to working memory and/or cognitive control.

Poster Session 1

MEG evidence for neural mechanisms in the reading of Chinese compounds
Chun-Hsien Hsu & Chia-Ying Lee (Academia Sinica)
kevinhsu@gate.sinica.edu.tw

Introduction. The majority of Chinese disyllabic words are compounds that have several sub-types. For example, subordinate words (SW) such as 汽車 /qi che/ gas-car (or an automobile) are right-headed, and the first morpheme is the modifier. Coordinative words (CW) such as 花草 /hua cao/ flower-grass (or plants) are headless, and their meanings are jointly derived from the two morphemes. For verb-object words (VO) such as 擔心 /dan xin/ carry-heart (or worry), the first and second morphemes express an action and an object, respectively. Behavioral studies have suggested that morphological structure plays a role in recognizing compounds. For example, studies using lexical decision tasks have demonstrated longer response latencies while reading CW and VO than while reading SW (Su, 1998; Zhang, 1997). This study aimed to further elucidate the mechanisms supporting the processing of compounds. Prior studies have suggested that the right anterior temporal lobe (RATL) and left posterior temporal lobe (LPTL) are involved in syntactic integration in reading syntactically anomalous sentences, and the left inferior frontal gyrus (LIFG) and left anterior temporal lobe (LATL) are involved in building syntactic structures (Grodzinsky & Friederici, 2006). In addition, the ventromedial prefrontal cortex (vmPFC) is involved in semantic composition (Bemis & Pylkkänen, 2011). On the basis of the assumption that linguistic processes governing the composition of words could direct the composition of morphemes (Halle & Marantz, 1993), we hypothesized that word types affect the activity in the regions of interest (ROIs) and reflect syntactic and semantic processes in recognizing compounds. To investigate the effects of semantic and syntactic complexity, we also compared brain responses while reading compounds to brain responses while reading monomorphemic-disyllabic words (MW) (e.g., 葡萄 /pu tao/ grapes). Finally, magnetoencephalography (MEG) data were acquired to examine the different stages of word recognition, such as the processing of lemma retrieval in early left temporal activities (Indefrey and Levelt, 2004).

Method. MEG data were recorded from 12 native Chinese-speaking adults in a lexical decision task. In the analyses, the targets were 200 disyllabic words, including MW, CW, SW, and VO. We used source analysis to trace the minimum norm activity in the ROIs. The effect of word types was evaluated using mixed-effects models, and we estimated fixed effects of three contrasts over the single-trial data, including one contrast between MW and compounds and two contrasts between compounds (CW vs. SW and SW vs. VO). Analyzing the single-trial data allowed us to incorporate potentially confounding factors into the analyses, including word frequency, number of strokes, and orthographic neighborhood size. The statistical significance of each contrast in each ROI was evaluated using non-parametric cluster-based analyses (Maris & Oostenveld, 2007).

Results. The earliest effect we observed was that LATL activity was larger in reading compounds than in reading MW at ~200 ms. The contrast between SW and CW showed opposite effects in two ROIs: reading SW revealed larger vmPFC activity than did reading CW at 419–483 ms, and reading CW revealed larger RATL activity than did reading SW at 418–475 ms. In addition, LATL activity was larger in reading VO than in reading SW at 367–413 ms.

Conclusion. This study indicates that different linguistic processes promote the reading of compounds. Early LATL activity reflects the building of compounds at ~200 ms. Semantic composition reflected by the vmPFC promotes the reading of SW. Finally, reading CW and VO involves more resources of syntactic integration.

What does left prefrontal cortex do for sentence production? Evidence from tDCS

Nazbanou Nozari (University of Pennsylvania), Jennifer Arnold (University of North Carolina at Chapel Hill) & Sharon L. Thompson-Schill (University of Pennsylvania)
Contact: nozari@sas.upenn.edu

Introduction. Anodal Transcranial Direct Current Stimulation (A-tDCS) has been shown to improve single-word production in both healthy speakers and aphasic patients. However, it remains unclear if A-tDCS changes production at the sentence level. We report a preliminary analysis of speed and accuracy of production under A-tDCS and sham, as participants described events online in a 'Moving Objects' paradigm (see below). If stimulation decreases the dual-tasking load (watching and utterance planning) we expect more efficient planning and fewer errors under A-tDCS. We would also expect this to affect choices of linguistic forms that require active linking to other discourse elements, namely pronouns/zeros, and perhaps to increase their use.

Methods. The data is reported from 6 participants in A-tDCS and 8 in sham. Participants were presented with 134 visual events involving one or two geometrical shapes on a grid and described them (e.g. "The green circle jumps over the purple trapezoid. The purple trapezoid moves two blocks to the left."). Data were analyzed using multilevel mixed models with random effects of participants and items.

Results. A-tDCS significantly decreased the number of sentences with errors ($t = 2.45$, $p = .014$). This decrease was observed across different types of words (See Figure 1). To investigate the possibility of speed-accuracy trade-off, we measured the speech-rate for a subset (20%) of the sentences, where no reliable difference was found between the two stimulation conditions ($t = .16$, $p = .87$). While in both conditions, combined pronouns and zeros were used 83% of the time the moving shape was the same as in the previous trial, the proportion of zeros (which link two utterances more explicitly than pronouns) was greater for A-tDCS (68%) than sham (44%).

Conclusions. A-tDCS reliably reduced the number of sentences with errors without slowing down speech, and increased the use of zero referential forms. Both of these point to more efficient use of resources for planning. These results suggest that tDCS is a useful tool for investigating the effect of cognitive control on sentence production, above and beyond what happens during lexical retrieval.
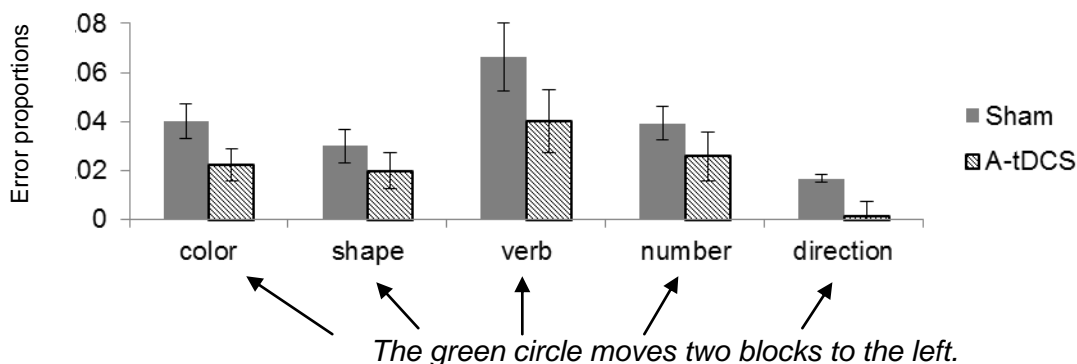

The green circle moves two blocks to the left.

Figure 1- Proportion of errors in A-tDCS and sham across different types of words.

Chinese aphasic patients' comprehension deficits with discourse-related constructions

Honglei Wang (Beihang University)
hongleiwang@buaa.edu.cn

Rationale. It has been observed that patients with Broca's aphasia are more impaired on discourse-related processes than on narrow syntax (Avrutin, 2006). For example, aphasic patients exhibit more difficulty comprehending pronouns (*him* or *her*), which involve discourse-level referent retrieval, than comprehending reflexives (*himself* or *herself*), which involve narrow syntax only (Grodzinsky, et al., 1993). Another construction involving discourse processes is with verbs that refer to the past tense. Bastiaanse, et al. (2011) report that patients showed more difficulty with verbs that refer to the past than those that refer to the present in tasks that test comprehension of verbs in different languages. They hypothesize that in comprehending verbs that refer to the past, patients need to establish a discourse relation between the speech time and the event time and therefore experience more difficulty than with verbs that refer to the present.

Prediction. The hypothesis by Bastiaanse, et al. (2011) predicts that patients should be impaired on both verbs that refer to the past and pronominals that involve discourse processes. In order to test this hypothesis, we tested three aphasic patients and three healthy controls on comprehending these two kinds of constructions in Chinese, specifically verbs that refer to the past and null objects whose reference is determined by a discourse topic.

Experiments. (1) First, we tested their ability to comprehend verbs that refer to the past vs. the present by using Test for Assessing Reference of Time (TART) (Bastiaanse, Jonkers & Thompson, 2008). Subjects were presented with a pair of pictures, and a sentence with either *le* to indicate the past tense or *zai* to indicate the present tense was read aloud. They were asked to point to the picture that matches the sentence. The accuracy data of patients showed that the percentage of correct responses was significantly higher for sentences that refer to the present than that for sentences that refer to the past, Z= -4.21, $p < 0.05$. Healthy controls did not show deficits on either tense. **(2) Subsequently**, we tested the same subjects' comprehension of null pronominals in the following constructions:

(1)   Context sentence: Zhangsan, Lisi and Wangwu were colleagues at the same university.
      a.   Zhangsan [*e* shuo hen xihuan Lisi].
           Zhangsan   say   very   like   Lisi
           'Zhangsan said that [*e*=Zhangsan] liked Lisi very much.'
           Question: Who liked Lisi?
      b.   Zhangsan shuo [Lisi hen xihuan *e*].
           Zhangsan   say   Lisi   very   like
           'Zhangsan said that Lisi liked [*e*≠Zhangsan] very much.'
           Question: Who did Lisi like?

It is argued in Huang (1984) that the null subject in (1a) is controlled by the matrix subject *Zhangsan*, involving only syntactic computation, while the null object in (1b) is linked to a discourse topic extra-sententially. The accuracy data of patients showed that the percentage of correct answers to the follow-up question was significantly higher in (1a) than in (1b), Z= -3.12, $p < 0.05$. Healthy controls, however, reached almost 100 percent accuracy in both follow-up questions.

Summary. The results in the two experiments demonstrate that patients had more difficulty with both reference to the past tense and discourse-related null objects. This result supports the hypothesis that aphasic patients are impaired on their ability to handle discourse related processes.

References. Avrutin, S. (2006). Weak syntax. In Y. Grodzinsky & K. Amunts (Eds.), *Broca's Region* (pp. 49–62). New York: Oxford University Press. Bastiaanse, R., Bamyaci, E., Hsu, C.-J., Lee, J., Duman, T. Y., & Thompson, C. K. (2011). Time reference in agrammatic aphasia: A cross-linguistic study. *Journal of Neurolinguistics, 24*, 652–673. Bastiaanse, R., Jonkers, R., & Thompson, C. K. (2008). Test for Assessing Reference of Time (TART). University of Groningen. Grodzinsky, Y., Wexler, K., Chien, Y.-C., Marakovitz, S., & Solomon, J. (1993). The breakdown of binding relations. *Brain and Language, 45*, 396-422. Huang, C. T. J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry, 15*, 531-574.

Distinguishing two routes to silent meaning in the brain

E. Matthew Husband (University of Oxford) & Fernanda Ferreira (University of South Carolina)
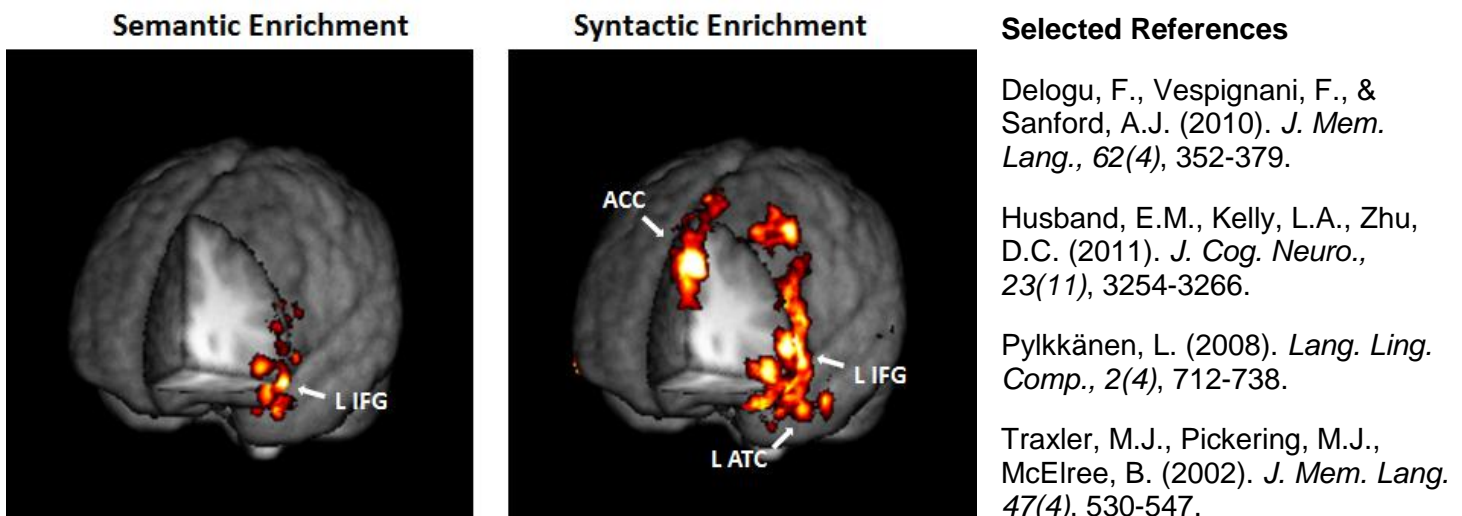matthew.husband@ling-phil.ox.ac.uk

Over the past decade, investigations of the processing of silent meaning have made inroads in our understanding of how sentence meanings are composed online using both behavioral (Delogu, Vespignani, & Sanford, 2010; McElree, et al., 2001; Traxler, Pickering, & McElree, 2002) and neuroimaging (Baggio, et al., 2010; Husband, Kelly, & Zhu, 2011; Kuperberg, et al., 2010; Pylkkänen & McElree, 2007) techniques. Native speaker intuitions suggest that sentences with silent meanings, like "*The reporter (1) began/(2) needed the article*", assert an implicit meaning (e.g. "*to read/to write/etc.*") which must be inferred and incorporated into the semantic representation of the sentence for successful comprehension. While these inferences appear to be similar, different computations are thought to derive them: (1) requires semantic enrichment, (2) requires syntactic enrichment (Pylkkänen, 2008). These computational differences may recruit different brain regions as semantic processing is thought to recruit left inferior frontal gyrus (LIFG) and left angular gyrus (LAG) while syntactic processing is thought to recruit LIFG and left anterior temporal cortex (LATC) (Lau, Phillips, & Poeppel, 2008).

To investigate this possibility, we conducted an event-related fMRI study contrasting sentences requiring semantic enrichment (1) or syntactic enrichment (2) with unenriched control sentences (*The reporter wrote the article*) and implausible sentences (*The reporter annoyed the article*). Thirteen adults read 336 sentences (84/condition) using word-by-word rapid serial visual presentation in four 8min 37sec blocks. Acceptability was judged after each sentence. fMRI data were acquired with echo planar imaging on a Siemens 3T scanner (8 channel head coil, 36 slices, 35msec TE, 2130msec TR, 90° flip angle, 208mm FOV, 64×64 matrix). fMRI preprocessing/analyses were conducted in FSL.

We report that semantic enrichment (vs. control) sentences elicited increased activity in LIFG (but not LAG). Syntactic enrichment (vs. control) sentences elicited increased activity in LIFG, LATC, and the anterior cingulate cortex (ACC). These results suggest that different neural circuits are required to process computationally different silent meanings. While semantic and syntactic enrichment both recruit LIFG, syntactic enrichment further recruits LATC, supporting syntactic computation, and ACC, which may aid in detection of syntactic requirements.

Figure 1. Differential activation for semantic and syntactic enrichment conditions minus control contrasts in the whole brain analysis. *Z* statistic images are thresholded using clusters determined by *Z* > 2.3 and a (corrected) cluster significance threshold of *p* = .05.

Selected References

Delogu, F., Vespignani, F., & Sanford, A.J. (2010). *J. Mem. Lang., 62(4)*, 352-379.

Husband, E.M., Kelly, L.A., Zhu, D.C. (2011). *J. Cog. Neuro., 23(11)*, 3254-3266.

Pylkkänen, L. (2008). *Lang. Ling. Comp., 2(4)*, 712-738.

Traxler, M.J., Pickering, M.J., McElree, B. (2002). *J. Mem. Lang. 47(4)*, 530-547.

Events along the garden path: A reduced N400 and a P600 in semantically reversible discourse

Gina Kuperberg, MD, PhD & Kristina Fanucci (Tufts University and Massachusetts General Hospital)
kuperber@nmr.mgh.harvard.edu

Introduction and Methods. We store vast amounts of knowledge within semantic memory, including information about familiar events, states and facts, as well as their likely relationships. Here we use event-related potentials (ERPs) to ask whether we are able to draw upon this type of knowledge to predict upcoming events during word-by-word sentence comprehension. ERPs were measured on target words as participants read four types of two-clause sentences (examples below): **(1)** *Event-related coherent*, in which the second event was causally related to the first; **(2)** *Event-related incoherent*, in which the events were temporally related to one another but the discourse connector, 'because', rendered their causal relationship incoherent; **(3)** *Lexically-related incoherent*, in which the events were unrelated but the target word was semantically related to lexical item(s) in the first clause; **(4)** *Unrelated incoherent* in which the two events were completely unrelated. 180 quadruplets of two-clause sentences were constructed (narrowed down in rating studies from an initial set of 320). The *Event-related incoherent* and the *Lexically-related incoherent* sentences were rated as equally incoherent at the point of the critical word. The critical words in the *Event-related coherent*, *Event-related incoherent* and *Lexically-related incoherent* sentences were equally semantically related to the 'bag of words' in their contexts, as indexed using Latent Semantic Analysis. There were no significant differences in frequency, length or orthographic density between the critical words across the four conditions. In the ERP study, 20 participants viewed the four sentence types, which were counterbalanced across four lists, such that no individual participant saw the same context in more than one condition. The sentences were presented word-by-word (450ms; ISI:100ms) and participants judged whether or not they made sense.

Results and Discussion. (1) As expected, the N400 was smaller to critical words in the *Event-related coherent* sentences than in the *Unrelated incoherent* sentences. Of most interest, the N400 was also attenuated on critical words in the *Event-related incoherent* sentences, despite these words being lexically unpredictable and incoherent. This suggests that comprehenders are able to draw upon stored event relationships to predict upcoming events in discourse, and that these predictions can facilitate semantic processing of incoming words, ahead of their being fully integrated into their discourse context, even if a causal connector renders these specific words lexically unpredictable. **(2)** Importantly, no such attenuation of the N400 was seen on the critical words in the *Lexically-related incoherent* sentences. This shows that event predictions cannot be reduced to the activation of simple lexico-semantic relationships. **(3)** The target words in the *Event-related incoherent* sentences also evoked a larger P600 than in all other conditions. We suggest that this reflected additional neural processing that was triggered when comprehenders detected the event prediction error: the conflict between the event sequence that was predicted and the event sequence that was computed when the incoming word was fully integrated into its context using the causal connector 'because'.

Conclusions. (1) We use stored relationships between events, beyond simple lexical semantic relationships, to predict upcoming events during comprehension, even in non-lexically constraining contexts. **(2)** Event predictions can influence semantic processing of incoming words, ahead of these words being fully integrated into their context; **(3)** Event predictions can also influence later stages of processing: if full integration of an incoming word into its context disconfirms these predictions, and this error is detected, the parser commits to additional analysis or reanalysis of the input.

Examples. ERPs were measured on critical words (underlined in these examples).
Event-related coherent: Kristen had insecticide applied to her house because there were cockroaches in her kitchen.
Event-related incoherent: There were cockroaches in Kristen's kitchen because she had insecticide applied to her house.
Lexically-related incoherent: Kristen had insecticide applied to her house because it was brick and metal throughout.
Unrelated incoherent: Kristen had insecticide applied to her house because she had a meal with her friends.

Predicting meaning: What the brain tells us about the architecture of language comprehension
Gina R. Kuperberg MD PhD (Tufts University and Massachusetts General Hospital)
kuperber@nmr.mgh.harvard.edu

It is well established that we draw upon our real-world knowledge to predict upcoming events (Altmann & Kamide, 99), and that we incur costs when the combinatorial integration of incoming words produces implausible events (Rayner et al., 04). In this talk, I will argue that that the neurocognitive mechanisms engaged in generating and using event predictions to retrieve conceptual information associated with incoming words are quite distinct from those engaged when these event predictions are disconfirmed by full combinatorial analysis of the input. Drawing broad links with computational models conceptualizing language comprehension as an incremental process of belief updating, I will suggest that the engagement of these distinct neurocognitive systems allows for comprehension that is both highly efficient and highly flexible.

First, I will discuss the N400 and P600 event-related potentials (ERPs). The N400 reflects the retrieval of conceptual information associated with a word or group of words ('surprisal' for a set of conceptual features). During sentence and discourse comprehension, the N400 evoked by an incoming word is modulated by conceptual information that is activated by that word's context. However, the N400 does not directly reflect the process of combinatorially integrating this word into its context to generate new meaning. Nor does it reflect the assessment of propositional coherence in relation to real-world knowledge. To illustrate these points, I will discuss studies showing that, in an event-constraining context, the N400 to an incoming word can be fully attenuated when that word's conceptual features match features that are associated with the predicted event or event structure, even when full combinatorial integration of this word, using all semantic and syntactic constraints, disconfirms this prediction, i.e. yields an event representation that is highly incoherent. In fact, under these circumstances, the brain produces quite a different response—a later, posteriorly-distributed P600 effect (e.g. Kuperberg et al., 2003; Kolk et al. 2003; Kim & Osterhout 2005, Paczynski & Kuperberg, 2012; Chow & Phillips 2013). I will argue that this P600 reflects a neural response that is triggered when the comprehender detects the discrepancy between the predicted and computed event—the *event prediction error* ('surprisal' for an event).

In the second part of the talk, I will draw some (still tentative) links between this ERP literature and some relevant fMRI and MEG studies. I will discuss evidence that predictive conceptual facilitation, indexed by the N400, maps on to reduced activity within the left anterior temporal cortex (Lau, Gramfort, Hämäläinen & Kuperberg, under review) and/or the posterior lateral temporal cortex (Lau, Poeppel & Phillips, 08). I will suggest that the left anterior inferior frontal cortex is engaged when conceptual features activated by a context mismatch conceptual features of incoming words. And I will present data showing that the P600 response, triggered by the detection of event prediction errors, patterns with the recruitment of the left inferior parietal cortex and dorsolateral prefrontal cortices in fMRI studies (Kuperberg et al. 03, 08).

The advantages of predicting correctly have been extensively discussed (e.g. Federmeier, 07; Pickering & Garrod, 12). In the final part of this talk, I will argue that the costs incurred when we predict incorrectly are also crucial for successful and flexible comprehension. First, I will discuss the hypothesis that the neurocognitive processes reflected by the P600/fronto-parietal response can rescue us from interpretation errors in noisy environments (van de Meerendonk et al., 09; cf Levy, 11). Second, I will suggest that they can allow us to recover novel meanings in sentences that at first appear to make no sense (cf Sitnikova, Holcomb & Kuperberg, 08; DeGrauwe et al., 10). Finally, I will discuss evidence that the P600 can be modulated by the predictability of the general (Coulson et al. 98) and immediate (Kuperberg, Lau & Clegg, 11) experimental environment. I will link these findings with the theories that event prediction errors play an important role in enabling us to flexibly adjust our comprehension strategies in response to ever-changing task and environmental demands (cf. Jaeger & Snider, 13).

Language processing in schizophrenia: Top-down & bottom-up effects
Hugh Rabagliati, Nate Delaney-Busch, Jesse Snedeker & Gina Kuperberg (Harvard & Tufts University)
hugh.rabagliati@ed.ac.uk

Introduction. A robust body of work now suggests that sentence comprehension is driven by the immediate parallel integration of both 'bottom-up' sensory/lexical information and 'top-down' discourse/pragmatic cues [Altmann, 1999]. One possibility, then, is that very similar mechanisms are used to integrate both high-level and low-level information sources. Neuropsychiatric populations may provide a reason to question this conclusion. For instance, in schizophrenia, it has been argued that a selective deficit of *top-down* integration may explain many of the associated cognitive impairments, including impairments to language comprehension [Cohen & Servain-Schreiber, 1992]. If so, top-down and bottom-up effects on sentence processing may dissociate in schizophrenia.

Methods. Using a fully within-subjects battery of four visual-world experiments, we compared top-down and bottom-up processing in 19 healthy controls and 19 patients with schizophrenia, matched on age, SES and pre-morbid verbal IQ. We examined how bottom-up information and top-down information are used to resolve prepositional phrase attachment ambiguities (*tickle the frog with the feather,* heard in front of a Target animal [frog with feather], Target instrument [feather], Distracter animal, and Distracter instrument. Each cue was manipulated to push the PP to either to serve as an **Instrument (Ins)** phrase [tickle [the frog] with the feather] or to **Modify (Mod)** the NP [tickle [the frog with the feather]]. **Bottom-up cues. Expt1: Prosodic Contour**: Ins *tickle the frog...with the feather;* **Mod:** *tickle...the frog with the feather;* 4 trials/condition. **Expt2: Structural Bias of the Verb: Ins:** *tickle the frog with the feather*; **Mod:** *find the panda with the stick;* 8 trials/condition. **Top-down cues. Expt3: Prior Discourse**. A 2nd speaker asked either **Ins:** *What should we do to a frog? or* **Mod:** *Which frog should we use now?,* followed by the 1st speaker giving the target instruction, *tickle the frog with the feather*, 4 trials/condition. **Expt4: Referential Context Ins:** 1 or **Mod:** 2 potential frog referents, 8 trials/condition. In each experiment, we crossed Bias (**Ins/Mod**) with Group (patient/control). We examined (a) participants' offline cue integration through the actions they ultimately performed; and (b) their ability to integrate these cues during online comprehension, as indexed by how Bias affected gaze to the Target instrument over 1500ms after the PP noun's onset, with statistically reliable time-windows extracted using resampling [Maris & Oostenveld, 2007].

Results. (a) Actions Patients and controls showed similarly successful offline integration of both bottom-up and top-down cues. In all experiments, both groups used the Target instrument more often in **Ins** trials (all effects of Bias, $p<.05$; no Group*Bias interactions, all $p>.25$). **(b) Online processing.** *In the bottom-up experiments*, the groups showed similar gaze patterns. Overall, **Ins**-biased **Prosodic** cues resulted in more looks to the Target instrument than **Mod**- biased cues, although the effect was delayed in patients (Bias*Group interaction 500-900ms, follow-ups: effect of Bias in Controls 500-1500ms, and in Patients 700-1500ms, all $p<.01$). Similarly, hearing **Ins**-biased **Verbs** led both groups to look more at the Target instrument than **Mod**-biased verbs, with similar time courses (effect of Bias 100-1500ms, $p<.01$, interaction ns.). *In the top-down experiments*, however, qualitative differences emerged between the groups. We found a Bias*Group interaction (600-1000ms, $p<.01$) that arose because **Discourse** affected whether controls looked at the instrument (400-1100ms, $p<.01$), but did not reliably affect patients' fixation patterns at any individual time-point). **Referential Context** showed a marginal Bias*Group interaction ($p\sim.1$), due to marginal reversed effects in each Group.

Conclusions. Patients with schizophrenia use both top-down and bottom-up information in *offline* actions, but show specifically impaired *online* integration of top-down discourse cues, despite comparatively spared bottom-up prosodic and lexical integration. Beyond informing our understanding of language in schizophrenia, this dissociation suggests an important distinction in the mechanisms by which top-down and bottom-up information are integrated online.

Predicting the *foreseeable future*: MEG evidence for preactivation of predicted words

Tal Linzen, Joseph Fruchter, Masha Westerlund & Alec Marantz (New York University)
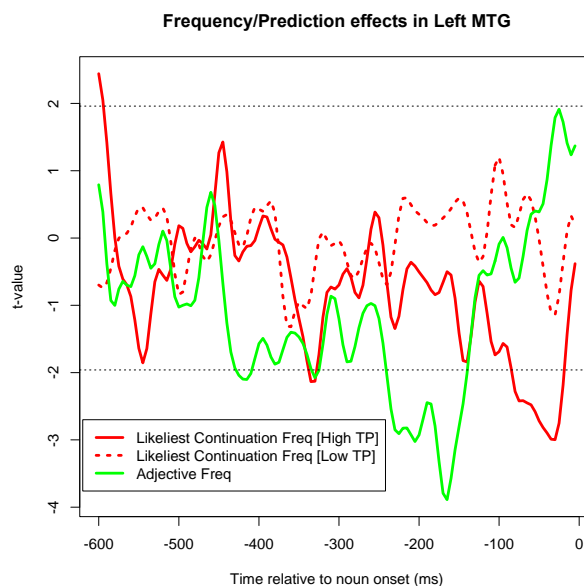linzen@nyu.edu

Introduction. Constraining context has a facilitatory effect on word processing: predictable words are read faster [1] and elicit less brain activity [2] than unpredictable ones. This facilitation can either be due to the pre-activation of the lexical entry before the word is encountered, or to increased ease of integrating the word into the existing context. Recent evidence suggests that words matched for ease of integration still show predictability effects, supporting the pre-activation account [3,4]. However, there is still little direct evidence for the formation of a word-specific prediction. The present study used MEG to detect lexical access of a predicted word before its presentation.

Methods. To maintain tight control over the point at which the prediction can be generated, we limited our materials to minimal adjective-noun phrases. We selected 475 such phrases varying continuously in transition probability (TP), as estimated from the Corpus of Contemporary American English. The phrases ranged from cases in which the adjective was highly predictive of a specific noun (*stainless steel*) to cases in which the adjective was not predictive of any particular noun (*important clue*). The adjective was presented for 300 ms, followed by a 300 ms blank screen that was then replaced with the noun. Participants performed a lexical decision on the noun. We obtained cortically-constrained minimum norm source estimates for MEG data from 16 subjects. We used an anatomically defined region of interest covering the left middle temporal gyrus (MTG), which has been implicated in lexical access [5].

Results. The presence of frequency effects can be taken to indicate lexical access. We found robust frequency effects in the left MTG both for the presented adjective and the presented noun, supporting the use of this region as an index of lexical access. We then used corpus transition probabilities to identify the noun most likely to follow each adjective, and analyzed whether MTG activity varied as a function of the frequency of this most likely continuation **before** the presentation of the following word. A single-trial correlational analysis revealed that this was indeed the case: the frequency of the most likely continuation modulated MTG activity at around 100 ms before noun presentation ($p < 0.05$).

Potential confounding variables, such as TP and adjective frequency, were controlled for. The effect held only for adjectives with a high TP continuation, such as *stainless* (usually followed by *steel*). This result presents a direct neural measure of lexical pre-activation, and argues against attributing predictability effects solely to ease of integration.

Figure. Correlation of stimulus variables with single-trial activity in the left MTG before the presentation of the noun. The adjective frequency effect reached significance around 200 ms before the presentation of the noun. The frequency of the likeliest continuation modulated activity 100 ms later, but only for predictive adjectives (adjectives which had a high TP continuation).

Frequency/Prediction effects in Left MTG

Likeliest Continuation Freq [High TP]
Likeliest Continuation Freq [Low TP]
Adjective Freq

t-value

Time relative to noun onset (ms)

References
[1] Ehrlich and Rayner, 1981. Journal of Verbal Learning and Verbal Behavior. [2] Kutas and Hillyard, 1984. Nature. [3] DeLong, Urbach and Kutas, 2005. Nature Neuroscience. [4] Dikker and Pylkkänen, in press. Brain and Language. [5] Lau, Phillips and Poeppel, 2008. Nature Reviews Neuroscience.

Interference in covert dependencies

Ming Xiang[1], Yanling Cui[2], Suiping Wang[2] (University of Chicago[1], South China Normal University[2])
mxiang@uchicago.edu

In *wh*-in-situ languages like Mandarin Chinese, for *wh*-questions, although the *wh*-phrase remains in a canonical argument position, syntactic theories generally posit that a dependency between the in-situ position and a clause-initial syntactic operator (at the highest [spect, CP] position) must nonetheless hold at Logical Form, rendering *wh*-in-situ languages and *wh*-fronted languages abstractly similar (Huang, 1982; Tsai, 1994). Processing wh-in-situ questions have also been shown to evoke extra cost compared to their non-*wh* parallel controls (Xiang et al. 2012). In current study, two self-paced-reading experiments showed that not only a long-distance dependency is being constructed online for *wh-in-situ* questions, but such process is also subject to memory interference that is known to affect the online processing of overt long-distance dependencies. For standard filler-gap dependencies, at the gap position the memory retrieval of an overt filler can trigger interference effect due to partial feature overlap between the retrieval cues and the intervening material (Lewis & Vasishth, 2005; Van Dyke & Lewis, 2003). For *wh*-in-situ questions, at the in-situ *wh*-phrase, the retrieval target is a certain [spec, CP] position. This provides the only structural cue for retrieval; no lexical-semantic, morpho-syntactic, or functional cues (e.g. subject or object) are available, since this position is not phonetically marked in any way. In Expt1 we show that such structural retrieval cues nevertheless can lead to interference.

Expt1 (n=53)-Embedded *wh*-questions rather than matrix *wh*-questions are used in this study. One short *wh*-condition is compared to three long ones, as in (1), in the Chinese word order:
(1)a.The reporter **wondered** the major punished **which officials**.   (Short)
b.The reporter **wondered** the major last-year summer punished **which officials**.   (Adverbial)
c.The reporter **wondered** the major led the city council punished **which officials**.   (Serial Verb)
d.The reporter **wondered** the major announced the city council punished **which officials**. (CP)
The three long *wh*-conditions were created by (i) adding a few adverbial modifiers (1b); (ii) adding an extra layer of serial verb structure (1c); or (iii) adding an extra layer of embedded CP structure (1d). The overall length of the three long conditions are matched, and for (1c) and (1d), the number of events introduced by the two sentences is also matched. A covert dependency needs to be established between the in-situ *wh*-phrase "*which officials*" and the [spec, CP] position located after "*wondered*". Critically, structurally speaking, only (1d) contains an additional intermediate [Spec, CP] position after the verb "*announced*". If pure structural features can induce interference, we expect (1d) to be more difficult than the other three. We also expect such interference effect to disappear in 4 additional declarative controls that share similar design features as (1) but do not involve covert dependencies. Self-paced-reading data (item n=64) showed that the in-situ *wh*-phrase in (1d) is read significantly slower than the other three conditions (ps<.01); there is no difference between the other three, suggesting longer linear distance alone does not induce extra processing cost. There is also no RT difference between control conditions. Consistent with this, (1d) is also rated the lowest on a 1-7 scale.

Expt 2 (n=48) shows that the effect in Expt 1 goes away if [spec, CP] is made irrelevant for retrieval. One type of overt *wh*-dependency in Chinese is topicalization, as shown in (2).
(2)a.   Which official, the major punished?   (Short)
   b.   Which official, the major last-year summer punished?   (Adverbial)
   c.   Which official, the major led the city council punished?   (Serial Verb)
   d.   Which official, the major announced the city council punished?   (Multi-CP)
At the matrix verb, to retrieve the topicalized *wh*-phrase, the set of retrieval cues would include the lexical features of the retrieval target, e.g. [+noun], [+topic], etc., but critically, not [spec, CP], since the *wh*-phrase is now in an extended [spec, Topic] position, not a regular [spec, CP] position. This would (i). render the intermediate [spec, CP] positions irrelevant for retrieval and interference; and (ii). make intervening nouns relevant for potential interference. Therefore, (2c) and (2d) become the targets for interference, due to additional intermediate nouns. Results showed longer RT on the matrix verb on (2c) and (2d), as well as lower rating on these two conditions, compared to (2a) and (2b) (ps<.05), with no difference within each pair.

Trainability and selective transferability of conflict resolution skills to parsing and non-parsing domains

Erika Hussey, Susan Teubner-Rhodes, Alan Mishler, Isaiah Harbison, & Jared Novick (University of Maryland)
ehussey@umd.edu

Conflict resolution (CR) refers to the executive function (EF) skill that enables individuals to resolve competing representations. During sentence processing, general-purpose CR skills may promote readers' ability to adjust parsing decisions when new input is incompatible with one's developing interpretation (i.e., ambiguity resolution). Indeed, a recent training study demonstrated that performance increases on a nonlinguistic CR task (n-back with lures) generalized to syntactic ambiguity resolution (e.g., A), but not unambiguous baselines that don't involve syntactic conflict (e.g., B; Novick et al., 2013). This was indexed by improved offline comprehension and real-time reanalysis: readers were more accurate on questions probing for misinterpretation *and* they spent less time regressing to early sentence regions after reading late disambiguating evidence (e.g., 'sparkled brightly' in A), where CR is hypothesized to engage.

Subjects in that study completed a battery of EF training tasks, while the comparison group (which enjoyed no benefits) completed no training, leaving open the possibility that better ambiguity resolution was due to general improvements in EF rather than specific improvements in CR ability. Here, we test whether CR training in particular transfers to ambiguity resolution by randomly assigning subjects to one of three alternative n-back training groups: High-Conflict, No-Conflict, or a static 3-back group. During n-back, subjects viewed single letters sequentially and indicated whether an item appeared n trials previously. N-level increased adaptively in the first two groups. Additionally, High-Conflict trainees encountered "lures"—recent items that appeared in non-n positions—compelling them to override a familiarity bias to correctly indicate that the item was not an n-back target. We predicted selective transfer from High-Conflict training to ambiguity resolution and no effect of No-Conflict or 3-back training to the same task, despite their memory-updating training component. Before and after training, all subjects completed non-parsing and parsing assessments with both high and low CR demands: (i) the Stroop task (with congruent and incongruent trials), (ii) an item recognition task (with and without interfering memoranda), and (iii) a reading task involving an ambiguity (A-B) and an object/subject relative clause manipulation (C-D). Crucially, object relative clauses (C) impose a memory demand compared to subject relative clause items (D; Fedorenko et al., 2006) but do not require the same degree of CR as (A).

As expected, although all training groups improved over the course of training and demonstrated better accuracy to questions probing for misinterpretations of (A), only High-Conflict trainees spent less time regressing to ($F(1,51)=4.68$, $p=.03$) and rereading earlier regions of ($Fs>4.76$, $ps<.03$) ambiguous sentences after encountering disambiguating evidence ('sparkled brightly'). The High-Conflict trainees demonstrated no changes in any reading time measure at posttest for sentences with low CR demands (B-D), irrespective of parsing difficulty, indicating the selectivity of their transfer effect to sentences requiring CR. Additionally, only the High-Conflict trainees improved on two non-parsing CR measures: smaller Stroop cost ($F(1,57)=4.06$, $p=.04$) and less susceptibility to memory interference at posttest ($F(1,58)=6.01$, $p=.01$), but no change in the non-interfering memory condition, verifying that CR was boosted for just this group of trainees. The findings provide support for the malleability of CR skills and suggest a critical (and perhaps causal) role for domain-general CR in sentence processing, perhaps independent of memory-updating demands.

- (A) While the thief hid the jewelry that was elegant and expensive sparkled brightly.
- (B) The jewelry that was elegant and expensive sparkled brightly while the thief hid.
- (C) The farmer who the expert questioned promoted the product at the fair.
- (D) The farmer who questioned the expert promoted the product at the fair.

References
Novick, J., Hussey, E., Teubner-Rhodes, S., Harbison, J., & Bunting, M. (2013). *LCP.*
Fedorenko, E., Gibson, E., & Rohde, D. (2006). *JML.*

The (un)automaticity of syntactic processing in language production: Structural priming is disrupted by verbal memory load

Iva Ivanova, Liane Wardlow Lane, Tamar Gollan & Victor Ferreira (University of California, San Diego)
iva.m.ivanova@gmail.com

Is syntactic processing in language production fully automatic (independent of other processes)? On one account, syntactic processing, at least in default circumstances, is automatic, unlike conceptualization processes, which demand processing resources (Bock, 1982). However, there is evidence that concurrent verbal memory load gives rise to more errors in subject-verb agreement (Fayol, Largy, & Lemaire, 1994; Hartsuiker & Barkhuysen, 2006), suggesting that at least certain aspects of syntactic processing demand resources shared with other (linguistic) processes.

In this study, we investigated the automaticity of syntactic processing in production with a structural priming paradigm. Specifically, we asked whether a concurrent verbal memory load would affect the structural priming effect. The answer to this question could also be informative about the structural priming phenomenon per se, which has been important in constraining models of language comprehension, production and acquisition (see Pickering & Ferreira, 2008). Additionally, this issue could have implications for models of language use in dialogue. This is because structural priming forms part of the core dynamics governing such use: Language production in a conversation is influenced by the linguistic input one receives (*interactive alignment*), facilitating the convergence on a situation model by the interlocutors, as well as promoting rapport between them (Pickering & Garrod, 2004).

We report a structural priming experiment in which 32 participants described dative-eliciting pictures (e.g., a waitress giving a banana to a sailor) after reading dative primes in a prepositional object (e.g., *The nun is throwing the apple to the pirate*) or a double object version (e.g., *The nun is throwing the pirate the apple*). During the experiment, participants 1) read a prime sentence on a computer monitor; 2) indicated with a button-press response whether this sentence matched a subsequently presented picture; 3) on Load trials, saw two words and were instructed to repeat them silently five times to remember them, for a total of 5 sec; on No Load trials, saw two rows of Xs for 5 sec (Load and No Load trials were mixed throughout the experiment); 4) described a target picture depicting a ditransitive event, with a verb written under it which they were instructed to use in their descriptions (this verb was always different from the prime verb); and 5) recalled the two words they had to remember on Load trials. There were 32 experimental items and 80 fillers.

Participants recalled both words correctly on 89% of all Load trials, and on 78% of the experimental Load trials. Considering only trials with correct memory load performance revealed a 25% priming effect in the No Load condition and a 12% priming effect in the Load condition. This pattern remained unchanged when all trials were included, when the subjects with worst load performance were excluded, and when a lax coding scheme of memory load performance was adopted.

In sum, structural influences of comprehended sentences on subsequent picture descriptions were diminished by concurrent verbal memory load. This result, together with recent evidence that conceptual information in production is mapped directly onto syntactic structure during a single stage (Cai, Pickering, & Branigan, 2012) suggest that syntactic processing is not fully automatic, possibly because verbal working memory is required for such conceptual-to-syntax mappings. Our finding potentially also implies that one might align less to their interlocutor (e.g., in a telephone conversation) if simultaneously scanning Facebook updates.

Domain-specific vs. domain-general mechanisms in language learning and processing

Ev Fedorenko (MIT)
evelina9@mit.edu

One of the oldest and most fiercely debated questions in cognitive science and neuroscience is whether our mind and brain are made up of component parts each specializing for a particular mental operation (e.g., Spurzheim, 1829; Broca, 1861; Brodmann, 1909; Fodor, 1983), or whether instead our brain is more equipotential with each brain region contributing to multiple cognitive processes (e.g., Flourens, 1824; Lashley, 1950). Language has long been at the epicenter of this debate. Chomsky (1957; also Pinker, 1994, among others) has strongly advocated domain-specificity of language. This position encountered skepticism and faced empirical challenges that highlighted the role of domain-general mechanisms in both language acquisition (e.g. Saffran et al., 1996) and processing (e.g., King & Just, 1991; Gibson, 1998; Gordon et al. 2002; Novick et al., 2005; Fedorenko et al., 2006, 2007). However, many discussions of domain-specificity vs. domain-generality of language in the last several decades have conflated conceptually independent issues. In this talk I will attempt to clarify two points: 1) Functional specialization does not imply innateness. 2) The existence of functionally specialized mechanisms for a mental process does not imply that domain-general mechanisms are not important for the process in question.

First, the existence of functionally specialized machinery in a mature organism need not imply that the relevant brain regions were specialized at birth (see e.g., Karmiloff-Smith, 1992; Elman et al., 1996, for discussion). Of course, some brain regions – plausibly evolved through natural selection – may be genetically "pre-programmed" to receive and process specific kinds of input (e.g., Cosmides & Tooby, 1994). However, functional specialization can also have experiential origins. For example, a region in the ventral visual stream known as the visual word-form area (e.g., Cohen et al., 2000) responds in a selective manner to letters in one's own script (e.g., Baker et al., 2007), and reading is too recent an invention to have a genetic origin.

A number of brain regions in the left frontal and temporal cortices have been shown to be functionally specialized for linguistic processing in both brain imaging (e.g., Fedorenko et al., 2011, 2012; Monti et al., 2009 2012) and neuropsychological (e.g., Varley & Siegal, 2000; Varley et al., 2005; Bek et al., 2010; Willems et al., 2011) investigations. On the surface, this evidence may be taken as supporting Chomsky's original position. However, an alternative (that I will argue for) is that the specialization of these brain regions develops via extensive experience with language (cf. Dehaene-Lambertz et al., 2006). Furthermore, it is possible that domain-specific knowledge structures (for example, our language knowledge) are acquired with domain-general statistical learning mechanisms (e.g., Saffran et al., 1996; Mintz et al., 2002; Samuelson, 2002 Perfors et al., 2012; see Saffran & Thiessen, 2007, for a review).

Second, the engagement of domain-general mechanisms in language is sometimes offered as an argument against domain-specificity of language (e.g., Amso & Blumstein, 2013). However, these arguments only apply to the notion of encapsulation, i.e., the idea that mechanisms that support the computations in a particular domain (e.g., language) do not interact with the rest of the brain (Fodor, 1983). But functionally-specific brain regions need not be encapsulated (e.g., Barrett & Kurzbann, 2006). Consequently, the existence of mechanisms dedicated to language processing is perfectly compatible with a role of highly domain-general mechanisms in language. Time permitting, I will discuss the nature of these mechanisms and their potential contributions to language comprehension and production.

Individual differences in verbal working memory predict reanalysis vs. integration difficulty in syntax-semantics conflict scenarios

Leif Oines, Albert Kim, Akira Miyake (University of Colorado Boulder)
leif.oines@colorado.edu

Psycholinguistic research has long sought to explain how combinatory processing is shaped by the brain's limited capacity to actively maintain information over time. Most research focuses on systematic processing costs for some linguistic structures over others (c.f. Just & Carpenter 1992, Gibson 1998, Caplan & Waters 1999, Macdonald & Christiansen 1996). Recent event-related brain potential (ERP) studies report that individuals with high vs. low working memory capacity showqualitatively different effects on the N400 and P600 ERP components in response to syntactically simple sentences. For instance, Nakano, Saron & Swaab (2010) investigated semantic P600 ERP effects and found that anomalous (animacy violation) verb-argument combinations elicited P600 in participants with high verbal working memory span scores, but N400 in lower-span individuals. Furthermore, Bornkessel, Fiebach & Friederici (2004) reported that garden path sentences elicited P600 in high-span individuals but N400 effects in low-span individuals. These findings raise the possibility that N400 and P600 reflect two different functional "options" that arise when the linguistic input is unexpected, with the choice between these options determined in part by individual differences in processing capacity. We recorded ERPs to "semantic attraction" anomalies which are thought to involve a syntax-semantics conflict, while simultaneously and systematically investigating how these individual ERPs varied as a function of multiple cognitive abilities: verbal working memory, spatial working memory, and language experience. Using multiple measures of each construct, we conducted regression analyses to specify which of them best accounted for individual differences in the brain responses to unexpected stimuli. We employed a sample of 60 subjects to obtain reliable correlations with individual capacity measurements. Across all subjects, semantic attraction anomalies elicited P600 effects, while no-attraction anomalies elicited N400 effects. Verbal WM predicted both individual subject P600 and N400 effect magnitudes, even when controlling for the effect of spatial WM and linguistic experience. As verbal WM increased, participants showed larger P600 and smaller N400 effects for both the attraction and the no-attraction violations. Spatial WM did not predict N400 or P600 effect magnitude after controlling for verbal WM, suggesting that the individual differences in ERPs are not accounted for by a general form of WM.

We suggest that syntax-semantics conflicts, which are pervasive in language processing, can lead to 1) an attempt to semantically integrate a syntactically licensed but implausible interpretation, reflected in N400 or 2) an attempt to restructure the sentence, reflected in P600, and that the latter restructuring response requires processing resources that are more available in high-capacity individuals than low-capacity individuals. Thus, verbal working memory capacity plays a critical role in the language processing system's management of linguistic information over time, and is at the heart of sometimes-qualitative inter-individual differences in language processing.

References
Bornkessel, I., Fiebach, C., Friederici A. (2004). On the cost of syntactic ambiguity in human language comprehension: An individual differences approach. *Cognitive Brain Research.*
Caplan, D. Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences.*
Gibson, E. (1998). Syntactic complexity: Locality of syntactic dependencies. *Cognition.*
Just, M.A., Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychol Rev.*
MacDonald, M.C., Christiansen, M.H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychol Rev.*
Nakano H., Saron C., Swaab T. Y. (2010). Speech and span: Working memory capacity impacts the use of animacy but not of world knowledge during spoken sentence comprehension. *Journal of Cognitive Neuroscience.*

Incremental parsing, gapping, and connectives

Masaya Yoshida (Northwestern University), Katy Carlson (Morehead State University), & Michael Walsh Dickey (University of Pittsburgh, VA Pittsburgh Healthcare System)
m-yoshida@northwestern.edu

When two clauses are connected, either by coordination (Frazier et al. 1984, Frazier et al. 2000, Knoeferle & Crocker 2009) or by subordination (Sturt et al. 2010), the processing of the second clause is facilitated if there is structural parallelism between them. But how strong is the preference for maximizing parallelism? We have found (Yoshida et al. CUNY 2011) that an ellipsis construction like gapping (1) is preferred by the parser over the temporarily possible structure where the second-clause subject is modified by the PP (*the lawyer near the telephone pole*).

1. [$_S$ The banker stood near the newsstand] and [$_S$ the lawyer ~~stood~~ near the telephone pole]

The gapping structure, though less common than a PP modifying a Noun, keeps the second conjunct fully parallel to the first conjunct. In this study, we compared the on-line processing of coordinated and subordinated clauses with or without parallelism. A grammatical constraint makes gapping possible in coordination but not subordination (Johnson 1994). We aimed to see whether maximization of parallelism would favor gapping on-line, and whether this preference would appear only in coordinated clauses.

The four conditions in (2) were tested in a word-by-word moving-window reading experiment (n = 40), with structural Parallelism (parallel vs. non-parallel) and Connective type (coordination vs. subordination) manipulated in a 2×2 factorial design.

2. a. The banker stood near the newspaper stand, and   [the lawyer near the telephone pole]…
   b. The banker tapped his feet,   and   [the lawyer near the telephone pole]…
   c. The banker stood near the newspaper stand, while [the lawyer near the telephone pole]…
   d. The banker tapped his feet,   while [the lawyer near the telephone pole]…
                                                                                        ...stood with crossed arms.

Crucially, the bracketed NP-PP sequence in each second clause is ambiguous: It can be analyzed as a gapping structure with a hidden verb ([NP [$_{VP}$ *~~stood~~ near* NP]]) or a simple NP with a PP modifier ([N [$_{PP}$ *near* NP]]). In this design, we expect an interaction of Parallelism and Connective at the following second verb if parallelism between two coordinated clauses leads the parser to expect gapping, since that verb disambiguates the sentence away from a gapping structure.

If structural parallelism facilitates the gapping analysis on-line, then we expect that the NP-PP string in (2a) should be preferentially analyzed as gapping. In (2b), however, parallelism does not hold between the clauses and the gapping analysis should not be facilitated. The PP should instead be analyzed as a modifier for the NP We therefore expect slower reading times at the second verb position in (2a) vs. (2b), because the second verb forces reanalysis to a non-gapping structure in (2a). In contrast, we do not expect a similar difference between (2c) and (2d), because in these two conditions, subordinating connectives make gapping ungrammatical.

We found that the verb stood in the second conjunct in (2a) was read significantly slower than in the other three conditions, i.e., a Parallelism × Connective interaction was found ($p < .05$). This result suggests that gapping is indeed preferred when there is structural parallelism, but only between two coordinated clauses. Note that the alternative noun modification structure is not only more economical than gapping, but also more frequent. However, the online positing of gapping is grammatically constrained: parallelism facilitates the ellipsis analysis only when it is grammatically licit, as in coordination.

Word order affects the time-course of sentence formulation in Tzeltal

Elisabeth Norcliffe[1], Agnieszka E. Konopka[1], Penelope Brown[1] & Stephen C. Levinson[1,2] ([1]Max Planck Institute for Psycholinguistics, [2]Radboud University Nijmegen)
elisabeth.norcliffe@mpi.nl

To produce a sentence, speakers must first generate a preverbal message and then map it onto a linguistic structure. How are these processes affected by word order? Eye tracking studies have shown that in English, sentence formulation may be lexically incremental [1]: speakers may begin their utterance having conceptually and linguistically encoded only one character (the subject), and may delay encoding of additional characters and the relationship between them until after they complete encoding of the subject character. Lexical incrementality is supported by English syntax because verbs are positioned sentence-medially: the event structure and the verb need not be planned in advance. In verb-initial languages by contrast, speakers *must* plan the verb prior to speech onset, and we hypothesize that this grammatical constraint may influence the extent to which speakers encode the event structure at the outset of formulation. We used eye tracking to study the time-course of sentence formulation in Tzeltal, a verb-initial Mayan language: specifically, we take advantage of word order variability in Tzeltal to test whether early verb production implies that speakers encode the relationship between the two characters earlier than when verbs are produced sentence-medially.

In Tzeltal, basic word order is VOS (1), although scrambled SVO ordering (2) is also possible:

    (1) ya   s-nuts   me'mut te   antse   (2) te antse ya s-nuts   me'mut
         ASP 3SG-chase chicken the woman   the woman ASP 3SG-chase chicken
         The woman is chasing a chicken   The woman is chasing a chicken

Thirty four native Tzeltal speakers participated in a picture-description eye-tracking study [1,2]. Target pictures, embedded in a list of unrelated fillers, showed two-character events that could be described with active or passive sentences. The distribution of responses on these trials (1145 sentences, 78% verb-initial and 22% verb-medial) was consistent with the overall preference for VOS word order [3]. Passives were produced slightly less frequently than actives (47% of the total dataset); the majority of passives had V-Oblique-S word order (89%).

We examined the distribution of fixations to the two characters over time for the different sentence types. The time-course of formulation for (active) VOS and SVO sentences showed that speakers were more likely to distribute their attention between the two characters in verb-initial sentences in a very wide time window (the first 2000 ms of a trial) than in verb-medial sentences. Specifically, formulation of SVO sentences proceeded as in English active sentences: speakers directed their gaze to the agent rapidly after picture onset and shifted attention to the patient approximately around speech onset [1], suggesting that encoding of the two characters was sequential. In contrast, formulation of active VOS sentences began with early fixations to the agent after picture onset, but this was rapidly followed by a sustained phase (400-2300 ms) in which fixations were distributed evenly between the agent and patient. This indicates earlier encoding of the patient and the action/verb than in SVO sentences. Finally, speakers redirected their gaze to the agent (the sentence-final subject) at around 2300 ms. A similar pattern was observed with passive verb-initial sentences: only after a phase of distributed attention between agent and patient did speakers direct their gaze to the patient (the sentence-final subject), at around 2300 ms.

The difference between the formulation of verb-initial and verb-medial sentences shows that speakers of Tzeltal adopt different strategies to encode these sentence types and, crucially, that the time course of encoding event participants may be mediated by the structural constraints of the language.

References
[1] Gleitman, January, Nappa, & Trueswell (2007). *JML, 57*, 544-569.
[2] Griffin, & Bock (2000). *Psych Sci, 11*, 274-279

The role of interactivity on cognitive alignment and decision making during dialogue

Moreno I. Coco[1], Rick Dale[2], Frank Keller[1] ([1]University of Edinburgh, [2]University of California, Merced)
mcoco@staffmail.ed.ac.uk

Dialogue is used cooperatively to guide joint action and achieve specific goals. Interlocutors build up a common ground of knowledge to achieve communicative success by aligning cognitive processes, both linguistic [1] and non-linguistic [2]. Within this framework, aligned cognitive processes should indicate optimal communication, which in turn should be associated with correct decisions. In the present study, we investigate the preconditions for successful alignment by testing the hypothesis that ***interactivity*** (bidirectional information flow) is crucial for alignment, and that optimal decision making in dialogue depends on it.

In 2 eye-tracking dialogue studies, we tested how decision processes depend on alignment under different conditions of interactivity.   In Experiment 1, communication was unidirectional, i.e., the speaker instructed the listener to perform a decision, but the listener could not give any feedback. Experiment 2 was bidirectional: the listener was allowed to give minimal feedback to the speaker, such as yes/no and backchannel utterances.   32 dyads of participants (16 per experiment) were engaged in a spot-the-difference task, where the speaker had to describe 100 photo-realistic scenes to the listener, who then had to decide whether he/she was viewing the same scene or not. Half of the time, the scene differed on a target object, which was either displaced in the scene (left, right), or absent in one of the scenes.

For the analysis, we divided the dyads into two groups according to their accuracy (high, low) and investigated the link between decision performance and gaze alignment. We first examined the speed-accuracy trade-off by looking at how reaction time correlated with decision performance. Then, we computed the alignment of eye-movement responses between speakers and listeners using cross-recurrence analysis and examined   the temporal shift at which listener gaze best matched the speaker gaze [2].

In Experiment 1, we found that longer reaction times correlated with better performance, but the accuracy of the listeners did not improve over the course of the experiment. In the absence of feedback, the listener needs to accumulate sufficient information to make accurate decisions. However, there is no improvement over time, possibly because common ground cannot be formed. In Experiment 2, instead, we found that reaction time did not play a role, and dyads became more accurate over trials. When some feedback is allowed, a common ground can be built, and decisions can be taken more confidently by the listeners. Concerning the recurrence of gaze, we found that in Experiment 1, low-performing dyads have a higher recurrence than high-performing dyads, whereas in Experiment 2, the opposite pattern was observed. Moreover, in both experiments, high-performing dyads display maximal recurrence at positive lags indicating the leading role of the speaker, while low dyads align better closer to lag 0, showing synchrony between speaker and listener. When there is interaction, aligned responses inform accurate decisions. Regardless of interactivity, accuracy in decision making emerges when the listener closely follows speaker's instructions.

Our data therefore suggests that interactivity is a precondition for optimal alignment and optimal decisions. In particular, when dialogue is not bidirectional, interlocutors have to balance *complementary* behaviors to compensate for the lack of interaction: high performance relies upon the listener decoupling from the speaker in their visual sampling of the scene. On the other hand, when minimal interactivity is allowed, common ground can be formed through alignment, and hence reliable decisions can follow from it.

References

[1] Brennan,S.E. & Clark, H.H., 1996, Conceptual pacts and lexical choice in conversation. Journal of Experimental Psychology: Learning, Memory and Cognition, *22*(6), 1482‑1493.
[2] Richardson, D.C. Dale, R. and Kirkham, N.Z., 2007, The art of conversation is coordination: common ground and the coupling of eye movements during dialogue, Psychological Science, 18, 407-413.

Motor plans and linguistic perspective in action sentences: A causal role in comprehension

Madeleine Beveridge[1], Daniel Casasanto[2], Roberto Bottini[2], & Martin Pickering[1] ([1]University of Edinburgh; [2]New School, New York)
m.e.l.beveridge@sms.ed.ac.uk

Recent research implicates some involvement of motor systems in language processing. However, the causal status of these motoric representations remains contested. We present a series of studies demonstrating a causal role for motor planning on sentence comprehension. Participants read a sentence describing a manual action in the first or third person, and choose from two photographs. On critical trials, both images depict the action described in the sentence: one shows a left hand performing the action and one shows a right hand performing the action. Pronoun (I, he) and image perspective (both photos internal, both photos external) are fully crossed. The dependent variable is the likelihood of choosing a right-handed image on critical trials.
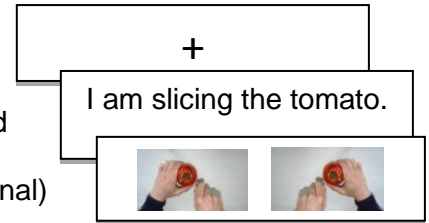
In Experiment 1 right-handed participants using their right hand were more likely to choose a right-handed image, while left- and right-handed participants using their left hands were both more likely to choose a left-handed image (all ps<.01). Importantly, the effect of response hand interacted with pronoun: all effects were stronger in first-versus third-person sentences (all ps <.05). There was no effect of image perspective. In Experiment 2 we adapted the paradigm, using a different coloured button box for each hand. Participants were shown a coloured screen to cue which box (and therefore hand) to respond with **after** reading the sentence. While reading the sentence, participants had no motor plan for a particular hand. There was no effect of response hand, or pronoun (all ps>.05). In Experiment 3, the coloured screen and button boxes were kept constant from Experiment 2, but participants responded with the same hand throughout a given block. Participants therefore knew, while reading the sentence, which hand would respond. Participants were more likely to choose a left-handed image when using their left hand, and a right-handed image when using their right hand (p<.01). The response-hand/ pronoun interaction was significant (p<.01), with a stronger effect in first- versus third-person sentences. The interaction between Expt. 2 and 3 was significant (p<.01).

Table 1. Percentage of trials where participants chose right-handed image in Expt. 1 (n=48)

	I am	He is
Right-hander/ right hand	58%	52%
Left-hander/ left hand	34%	38%
Right-hander/ left hand	39%	46%

Table 2. Percentage of trials where participants chose right-handed image in Expt. 2 (n = 16) and Expt. 3 (n = 16)

	I am	He is
Right hand response (expt. 2)	48%	54%
Left hand response (expt. 2)	53%	53%
Right hand response (expt. 3)	67%	56%
Left hand response (expt. 3)	41%	43%

Experiments 1-3 demonstrate that sentence-picture matching is affected by the hand you are planning to respond with, *at the point of reading the sentence*. If participants are prevented from forming a motor plan until after reading the sentence (Expt. 2), the subsequent plan has no effect on sentence interpretation. Moreover, the more strongly a comprehender views herself as the possible agent of the sentence, the stronger the effect of her motor plan on processing that sentence is. Whether a comprehender interprets a sentence as referring to herself is affected by subject (I/ he), but not image perspective (internal/ external). These results support a causal role for motor planning in sentence processing: when a motor plan is present, the nature of that plan interacts with the sentence context (in this case, subject pronoun) to affect sentence interpretation. This does not entail that all comprehension must be mediated by the motor system (participants in Expt. 2 still answered correctly on >90% of filler trials). Rather, we argue that when a motor plan is present during sentence comprehension, the resulting interpretation is different to one arrived at without activation of relevant motor systems, and dependent on the nature of that motor plan. Sentence comprehension is a flexible and interactive process which draws on all resources available – in this case, the intention to move a particular hand.

Perspective taking in online language processing
Xiaobei Zheng, Richard Breheny (University College London)
xiaobei.zheng.10@ucl.ac.uk

Research into the use of perspective information on line tends to adopt a binary representation framework, where one's own and common knowledge are represented separately (Keysar et al, 2000). In contrast, Horton and Gerrig (2005) propose that common ground is an emergent property of memory processes, leading to the single complex gradient representational framework (Brown-Schmidt & Hanna, 2011). According to the latter view, common ground information shared with a partner is salient to the extent that the partner provides reliable cues to common ground information. We present two studies supporting the latter view.

In two eye-tracking studies, participants were asked to move objects around a 3 by 3 grid by clicking on and dragging images. At the beginning of each item, participants placed three privileged objects (not known to the speaker) in the three grid positions that the speaker cannot see. As in similar visual-world studies (Keysar et al, 2000) test items ('Move the apple to the bottom middle') occurred when there was a competitor (Privileged condition) or an irrelevant object (Baseline condition) in their privileged ground. Filler items included location-based directions ('move the object in the top left…') as well as the more standard type-based instructions ('Move the apple…') used also in test items. In a homogeneous cue condition, all instructions made reference only to common ground locations. In a heterogeneous condition, location-based filler instructions made reference also to participants' privileged ground objects. In the first experiment participants were divided into either homogeneous group or heterogeneous group, while in the second experiment all the participants had two directors, one gave instructions for the homogeneous and one for the heterogeneous condition.

According to the single gradient framework, only common ground objects were associated to the director in the Homogeneous condition, but in the Heterogeneous condition the privileged ground is also associated with the director and thus he/she is not as reliable a cue to common ground objects, so performance should diminish. By contrast, if participants hold a separate representation of common ground location-based reference to privileged ground slots should not affect this.

Participants' eye movements were recorded and analyzed in two 300ms windows after a baseline region (-200ms to 200ms from the onset of target word). In experiment 1, there was a significant interaction between common ground and association conditions during 200 to 500ms, $F1(1,38)=16.259$, $p<0.01$, $F2(1,14)=12.186$, $p<0.01$, and a trend in 500 to 800ms, $F1(1,38)=3.131$, $p=0.085$, $F2(1,14)=3.663$, $p=0.076$. In the Privileged condition participants had larger target advantage scores in Homogeneous condition than Heterogeneous condition during 200 to 500 ms, $F1(1,38)=9.296$, $p<0.01$, $F2(1,14)=12.340$, $p<0.01$, and also 500 to 800 ms, $F1(1,38)=3.900$, $p=0.056$, $F2(1,38)=5.987$, $p<0.05$. In experiment 2 we found a similar interaction in critical windows, though items analysis only trend due to small number of items: 200 to 500ms, $F1(1,30)=4.547$, $p<0.05$, $F2(1,6)=4.151$, $p=0.088$, and 500 to 800ms, $F1(1,30)=4.396$, $p<0.05$, $F2(1,6)=3.101$, $p=0.129$. In the Privileged condition participants had larger target advantage scores in Homogeneous condition than in Heterogeneous condition during 200 to 500 ms, $F1(1,30)=5.702$, $p<0.05$, $F2(1,6)=5.069$, $p=0.065$, and during 500 to 800 ms, $F1(1,30)=4.693$, $p<0.05$, $F2(1,6)=4.172$, $p=0.087$.

The results in both experiments indicated that in the Heterogeneous condition participants got larger interference from their privileged ground compared to the Homogeneous condition, even though the common and privileged ground objects were all the same in both conditions. The results showed that the director is a better cue for common ground information in the Homogeneous condition than the Heterogeneous condition, and support the complex gradient representational framework.

References
Keysar, Barr, Balin,& Brauner, (2000). Psych Science, 11(1), 32-38.
Horton & Gerrig, (2005). Discourse Processes, 40(1), 1-35.
Brown-Schmidt, & Hanna, (2011). Dialogue and Discourse 2, 23.

Discourse-driven biases in native- vs non-native speakers' coreference processing

Theres Grüter[1], Hannah Rohde[2], & Amy J. Schafer[1] ([1]University of Hawai'i; [2]University of Edinburgh)
theres@hawaii.edu

Native-language processing involves not only information *integration*, but also *anticipation* (e.g., AltmannKamide99). At the discourse level, the ability to anticipate upcoming outcomes has been shown to reflect a variety of local contextual cues. For example, expectations about upcoming patterns of coreference and coherence show sensitivity to semantic properties of the current sentence, such as the event structure conveyed by the verb (KehlerEtAl08). Given that non-native processing is generally subject to greater resource demands (Kilborn92), a question emerges as to whether non-native speakers' coreference processing will show less semantically-driven anticipation. If so, non-native speakers—even those who show mastery of the linguistic cues that mark event structure—may rely more on superficial entity salience, such as a bias for coreference to recently mentioned entities (as L1 children do, see KehlerEtAl11).

A story-continuation task (adapted from RohdeEtAl06) was completed by 20 advanced L2 learners of English (12 Japanese, 8 Korean) and 24 native speakers; they wrote continuations following a context sentence that described a transfer-of-possession event (see (1)). A 2x2 design varied *aspect* in the context sentence (perfective/imperfective) and *prompt type* in the continuation (pronoun/free). Trained judges identified the intended referent of the subject of the continuation, coding for SOURCE (of the context sentence; see (2)), GOAL (3), ambiguous, or other. Previous work shows that native speakers' coreference expectations on this task are modulated by verbal aspect: Transfer-of-possession events yield more Goal-reference when marked by perfective than imperfective aspect. This effect, observed in English (RohdeEtAl06), Japanese (Ueno&Kehler2010) and Korean (Kim2012), has been tied to *end-state salience*: perfectives describe completed events compatible with end-state focus, whereas imperfectives describe ongoing events with no salient end-state. Event structure plays a role in native speakers' coherence expectations, which in turn guide coreference—e.g., Narration coherence relations, which tell what happened next, are expected in contexts with a salient end-state and favor reference to the end-state entity (here, Goal), while Explanation and Elaboration relations, which explain why or how an event occurred, are expected following ongoing events and favor reference to the individual associated with the event onset (here, Source). Pronoun interpretation biases observed in the story-continuation task thus reflect the deployment of a chain of pragmatic biases and expectations. Importantly, an independent task verified our L2 group had native-like knowledge of the ongoing vs. completed function of aspect. Nevertheless, if non-native speakers are less able to engage in predictive discourse-level processing, the L2 group should show a reduced effect of aspect on coreference biases.

Results (see Fig.1) show differential performance by group (L1, L2), with the predicted aspect-by-group interaction ($F1_{(1,42)}$=7.34, $p<0.01$; $F2_{(1,9)}$=8.66, $p<0.05$) driven by the influence of aspect on the proportion of Goal/Source reference (for prompts of both types) by native but not non-native speakers. This suggests that L2ers have a reduced ability to use event structure to generate discourse expectations. Moreover, L2ers showed an overall Goal bias, suggesting that recency plays a stronger role in L2 than native processing. Both groups demonstrated similar Goal/Source coreference biases across coherence relations, but in keeping with their Goal bias, L2ers favored Goal-biased relations. This suggests that L2ers may have native-like knowledge of event structure, but instead of using coherence expectations to guide coreference, their recency-related Goal-bias drives a subsequent choice of coherence relation.

Fig.1. % Source-reference by Group (L1,L2) and Prompt Type (pronoun, free)
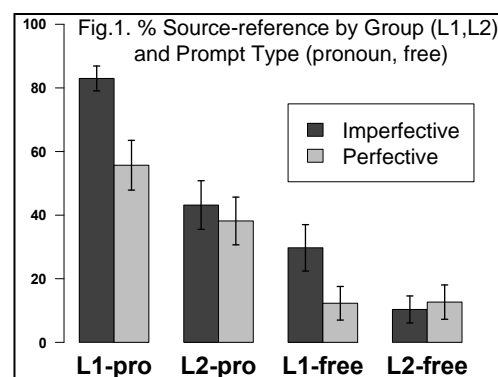
(1)   Experimental context sentence: Patrick gave/was giving a towel to Ron. (He) __
(2)   SOURCE-continuation for (1), i.e., *He* = Patrick: *He made sure to give him a clean dry one.*
(3)   GOAL-continuation for (1), i.e., *He* = Ron: *He took it and said "Thank you."*

Poster Session 2

Online processing of English garden-path sentences by L2 learners: A visual world study

Lucia Pozzan & John Trueswell (University of Pennsylvania)
lpozzan@sas.upenn.edu

Introduction. An important issue in psycholinguistics concerns the extent to which processing strategies change across development and are influenced by the maturation of cognitive and linguistic abilities. It is well established that English-speaking children differ from adult native speakers in their ability to recover from initial misinterpretations.[1] While children's difficulties recovering from garden-path sentences have been linked to immature cognitive control, a growing body of research in second language (L2) acquisition shows that *adult* learners experience similar difficulties[2,3] suggesting that language experience and proficiency might be additional factors affecting learners' ability to recover from strong garden-paths. Here, we investigate how adult L2 learners of English interpret temporarily ambiguous sentences, using the same visual world paradigm and analogous task and materials as those used with child learners of English. We examine the extent to which L2 learners differ from monolinguals in how they interpret and revise temporally ambiguous sentences, and in their ability to use referential cues to guide initial interpretations. Participants heard temporally ambiguous and unambiguous sentences (e.g., *Put the frog (that's) on the napkin onto the box*). The contexts included either one or two referents (e.g., frogs), supporting either a goal or a modifier interpretation of the first prepositional phrase (PP), respectively. L1-Italian learners were chosen because Italian presents the same temporal ambiguity as English; any differences between the L2 learners and the English native speakers should thus reflect added difficulties due to processing a weaker, non-dominant language.

Method. <u>Task</u>: Participants acted out spoken English commands by moving objects on a computer screen while being eye-tracked. <u>Materials</u>: 24 experimental sentences, and 102 fillers. <u>Participants</u>: 33 L1-Italian L2-English adults (mean MTELP score=35/45, SD=6), 24 adult native English speakers.

Results and discussion. Proportions of looks to the incorrect-destination (e.g., the empty napkin) after the onset of the first PP appear in Figures 1 and 2 for L2 and native speakers, respectively. Multi-level mixed logit modeling show significantly more looks to incorrect-destinations for ambiguous vs. unambiguous sentences ($p<.05$), 1-Referent vs. 2-Referent contexts ($p<.05$) and L2 learners vs. native speakers ($p<.05$). Moreover, a significant ambiguity by language group ($p<.05$) interaction emerged: while L2 learners, overall, showed higher proportions of looks to the incorrect destination, they were especially affected by the Ambiguity manipulation. A significant ambiguity by referential context interaction emerged ($p<.05$) in the absence of a three way interaction with language group ($p>.05$), indicating that both groups experienced stronger garden-paths in Ambiguous 1-Referent than in 2-Referent contexts. A similar pattern of results emerges from the offline actions (not shown here). Taken together, these results indicate that interpreting and revising temporarily ambiguous sentences is a hard task for language learners, irrespective of age of acquisition. On the other hand, in line with a number of findings in the L1 and L2 literature, these data also suggest that adult learners are more successful than children at using referential pragmatic-discourse information to drive interpretations. Ongoing research investigates the nature of this difference and the extent to which it is modulated by language background and proficiency.

Figure 1: Adult L2 Learners

Figure 2: Adult Native

References

[1] Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). *Cognition.* [2]Williams, J. N., Möbius, P., & Kim, C. (2001). *Applied Psycholinguistics.* [3]Roberts, L., & Felser, C. (2011). *Applied Psycholinguistics.*

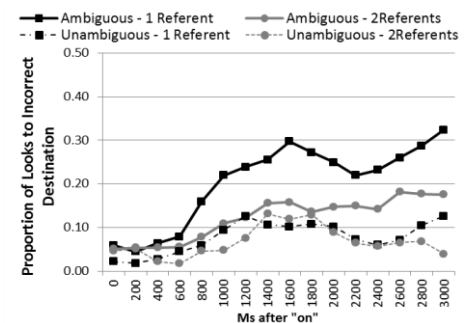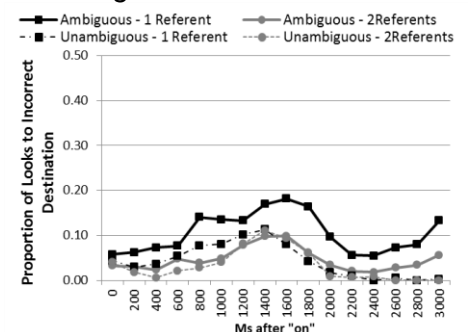**Word order and interference in online gap-filling by bilinguals**
Irina A. Sekerina (College of Staten Island and the Graduate Center, CUNY)
irina.sekerina@csi.cuny.edu

Data from a variety of languages and participant groups are critical in advancing theories of processing of syntactic dependencies in a search for an explanation why young children, people with aphasia, and bilingual speakers make comprehension errors in sentences with noncanonical object *wh*-questions, like (1a-b). The Visual World eye-tracking paradigm (VWP) allows us to pinpoint the locus of this difficulty and to propose a unified explanation.

Dickey et al. (2007) showed that when English-speaking aphasic patients were able to process sentences like (1b) correctly (70%); they showed the same automatic association of the *wh*-word with its trace as controls (i.e., increased looks to the object at the verb). When aphasics made errors, it was because of competition from the subject (i.e., increased looks to it) in the post-verb regions. Dickey et al. took these results as evidence against impaired-representation and slowed-processing accounts of aphasic comprehension deficits.

We present the results of the VWP experiment on Object *wh*-questions in Russian with control (*N*=8) and bilingual heritage (*N*=24) participants in which case marking and flexible verb position speed up association of the *wh*-word with its trace. Participants listened to short stories and answered a *wh*-question by clicking on 1 out of 4 pictures while their eye movements were recorded. Materials were 40 subject- (1c-d) and object- (1a-b) *wh*-questions with preferred and scrambled word orders, with the *wh*-words unambiguously marked for case (NOM for subj, ACC for obj). Overall bilinguals' comprehension was high (90% accuracy), but there was a Type of Question x Word Order interaction (no difference for 1d and b; 9.5% difference for 1c and a).

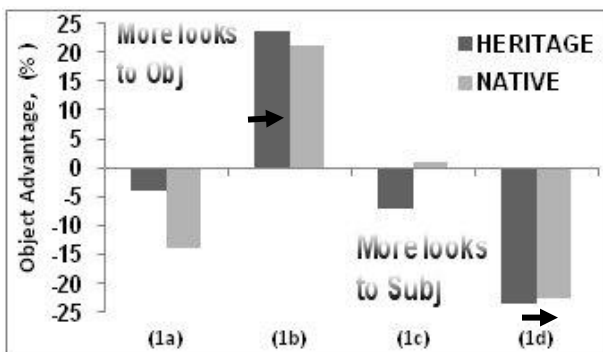| (1)		**Word Order**				**Accuracy (Bilingual)**	
Object	a.	**Wh-ACC V S** (preferred):	*Kogo*	*spas*	*zajac*	*v jame?*	**(86.5%)**
		(Russian≠English)	who_ACC	saved	rabbit_NOM	in the hole	
	b.	**Wh-ACC S V** (scrambled):	*Kogo*	*zajac*	*spas*	*v jame?*	**(91.7%)**
		(Russian=English)	'Who₁ did the rabbit save __₁ in the hole?				
Subject	c.	**Wh-NOM V S** (preferred):	*Kto*	*spas*	*kozla*	*v jame?*	**(95.8%)**
		(Russian=English)	who_NOM	saved	goat_ACC	in the hole	
	d.	**Wh-NOM S V** (scrambled):	*Kto*	*kozla*	*spas*	*v jame?*	**(91.1%)**
		(Russian≠English)	'Who saved the goat in the hole?'				

Fig. 1. Object advantage in looks at the Verb region

Eye movements at the Verb region (Fig. 1) showed that in the traceless (1c-d), there was no difference in looks (Subj=Obj) and no interference from the intervening Obj for either group. In the filler-gap (1a-b), there was an interaction: both control and bilinguals showed strong and automatic association of the *wh*-word with the trace in (1b) where word order (SV) in Russian and English is the same, but only bilinguals experienced more Subj/Obj competition in Russian-specific VS (1a).

This new evidence supports the Cue-Based Retrieval Theory (Van Dyke & McElree, 2006) that argues for interference-based competition as the source of difficulty with comprehension of *Wh*-movement structures in aphasics (Dickey et al.), but it comes from bilinguals who do not have disturbance in their syntactic competence. It enables us to propose a unified explanation of how a variety of cognitive states (e.g., language breakdown in aphasia and cognitive consequences of bilingualism) can affect retrieval efficiency.

Poster Session 2

Syntactic constraints in the processing of wh-movement by L2 learners
Adrienne Johnson, Alonso Canales, Rob Fiorentino & Alison Gabriele (University of Kansas)
johnsonam@ku.edu

Introduction. The current study investigates whether L1 Spanish and L1 Korean learners of English are able to utilize knowledge of abstract syntactic constraints in the on-line processing of wh-dependencies, similar to native speakers of English. Previous studies have shown that English native speakers are both incremental and accurate in their processing of sentences involving wh- filler-gap dependencies (Stowe, 1986). The Shallow Structure Hypothesis (Clahsen and Felser, 2006) predicts that second language (L2) learners are not guided by the same syntactic constraints during on-line processing. Additionally, the Failed Functional Feature Hypothesis (Hawkins and Chan, 1997) predicts that L1 Korean learners of English will be unable to acquire the syntactic properties of wh-movement because Korean is a wh-in-situ language and learners are limited to the syntactic features present in their L1 (e.g. +wh).

Method. The current study used a moving window self-paced reading paradigm (Just, Carpenter & Wooley, 1982) to examine whether L2 processing of wh-dependencies is incremental (Experiment 1) and guided by syntactic constraints on movement (Experiment 2). If processing is incremental, upon encountering a wh-filler (e.g. *who*), the parser will attempt to resolve the wh-dependency by positing a gap at each potential gap position. Thus, as measured in Experiment 1, the parser will experience processing difficulty as evidenced by a reading time slowdown or *filled gap effect* when a potential gap position is already filled with lexical material (e.g. direct object *Tom*) in the wh-extraction sentence (b) as compared to the declarative sentence (a).

a. The instructor wondered *if* Chris will film **Tom** with Susan at the reception.
b. The instructor wondered *who* Chris will film **Tom** with _____ at the reception.

Experiment 2 examined whether learners use knowledge of syntactic constraints in order to avoid positing gaps in a position in which a gap is not licensed syntactically, such as within an relative clause island. Despite the presence of a gap-licensing verb (e.g. *married*), no slowdown should occur at *Tyler* in (d) as compared to the same position in (c) because extraction out of a relative clause island is not permitted. The Shallow Structure Hypothesis predicts that L2 learners will attempt to illicitly posit a gap following the verb within the relative clause island.

c. My father asked *if* [the actress that married **Tyler** last summer] kissed the director . . .
d. My father asked *who* [the actress that married **Tyler** last summer] kissed _____ . . .

Results. Reading times were measured at the direct object position (e.g. *Tom/Tyler*) and the spillover regions in the wh-extraction conditions (b, d) and compared to their respective declarative conditions (a, c). Similar to native speakers (N=48), advanced Spanish (N=54) and advanced Korean (N=48) learners of English show evidence of incremental parsing in the form of a reading time slowdown (i.e. filled gap effect) in the spillover region in the wh-extraction condition in Experiment 1 ($p < 0.05$). Also similar to native speakers, advanced Spanish and Korean learners of English avoid positing gaps in positions prohibited by syntactic islands as evidenced by an absence of filled gap effects within the relative clause islands in Experiment 2.

Discussion. Our findings suggest that L2 learners are able to use the same abstract syntactic information in on-line processing as is used by native speakers, contra the claims of the Shallow Structure Hypothesis (Clahsen and Felser, 2006). Furthermore, our study provides evidence that L2 learners have access to such abstract syntactic information even when their native language (e.g. Korean) does not instantiate the same properties.

Unfolding an event differently: An ERP study on L1 and L2 processing of grammatical aspect
Shengyan Long, Manami Sato, & Hiromu Sakai (Hiroshima University)
longshengyan@hiroshima-u.ac.jp

Introduction: How two languages are organized and skillfully processed within an advanced L2 learner? While recent ERP studies have shown that L1 and L2 speakers elicited different ERP components when processing syntactic violations (Hahne, Mueller, & Clahsen, 2006), others have shown that L2 can be processed in an L1-like manner when L1 and L2 share the similar syntactic features (e.g., Gender agreement system, Sabourin & Haverkort, 2003; Sabourin & Stowe, 2008). Nevertheless, types of target languages as well as syntactic features that have been examined are quit limited (e.g., English person/number agreement and subjacency constraint, Ojima et al., 2005, Weber-Fox & Neville, 1996; German phrase structure, Hahne and Friederici, 2001, for instance). In this study, we report two Japanese ERP experiments that examine how L1 and L2 speakers unfold an event described in sentences using aspectual information. We explore the questions of (1) how Japanese native speakers and Chinese advanced learners of Japanese process aspectual information and (2) whether or not L2 can be processed in an L1-like manner even when morphosyntactic properties of L1 (Chinese) and L2 (Japanese) are different.

Experiment 1. In Experiment 1, twenty-six Japanese native speakers read Japanese sentences such as (1), where the aspectual property of the adverb "*mada* (until)" either matched or mismatched with the aspectual marker, "*-teiru* (progressive)" or "*-ta* (perfective)", respectively. Compared with correct sentences, the aspect violations elicited a left anterior negativity (LAN, 300-500 ms) for native Japanese speakers. This suggests that L1 processing of grammatical aspect is not semantically driven, but is morphosyntactically operated (Friederici, 2002).

Experiment 2. In Experiment 2, sixteen Chinese speakers with high Japanese proficiency read the same Japanese sentences as in Experiment 1. L2 learners showed a 300-500ms posterior and left central negativity. This is distinct from ERP responses observed in L1 speakers, but similar in the timing and distribution of the ERPs that were elicited by Chinese speakers when processing violations of grammatical aspect in Chinese (Zhang & Zhang, 2008).

Discussion. Although L2 learners performed sentence judgments correctly and detected the aspect violations like native speakers, distinct patterns of neural activities were observed in Experiments 1 and 2. This suggests that L1 and L2 learners of Japanese may involve different cognitive mechanisms when they process aspectual information that denotes how an event temporally unfolds. Moreover, L2 learners cannot process L2 sentences in the same manner as native speakers if their L1 and L2 have different morphosyntactic systems.

(1) a. Match: ototo-wa   mada   sara-wo   arat-teiru   youdesu.
              brother-NOM still-IMPERF a plate-ACC wash-PROG look like
              'It looks like that my brother is *still washing* the dishes.'
   b. Mismatch: ototo-wa   mada   sara-wo   *arat-ta   youdesu.
                          wash-PERF
                 '*It looks like that my brother *still has washed* the dishes.'

Predictive use of case marking during sentence comprehension: An eye-tracking study of Turkish-speaking children (and adults)

Duygu Özge[1], Aylin Küntay[2], & Jesse Snedeker[1] ([1]Harvard University; [2] Koç University)
dozge@wjh.harvard.edu

Adult language comprehension is rapid and flexible: in head-initial languages like English, upcoming arguments are predicted on the basis of verb information (Altmann & Kamide, 1999), while in head-final languages like German (Kamide, Scheepers & Altmann, 2003) and Japanese (Kamide, Altmann & Haywood, 2003) case marking is used predictively. How does this system develop? Prior studies provide ample evidence for the predictive use of verb information in children learning languages like English (Snedeker & Trueswell, 2004; Fernald, 2004). But we do not know whether children learning head-final languages use case markers for predictive processing. Choi and Trueswell found a parsing preference which suggests that Korean-speaking children are sensitive to case marker frequency, but they did not manipulate the case marker (Choi & Trueswell, 2010). German-speaking children fail to use case markers to interpret OVS constructions until age six or seven (Dittmar, Abbot-Smith, Lieven & Tomasello, 2008), leading some theorists to propose that the integration of case during comprehension is subserved by a late maturing neural system (Knoll, Obleser, Schipke, Friederici & Brauer, 2012; Friederici, 2012).

The present study explores whether Turkish-speaking children integrate case marking to make predictions about upcoming arguments and whether they can do this independent of the verb information. Turkish is a verb-final language with regular case morphology and variable word order. This study tests whether nominative (unmarked case) and accusative case (-i) on NP1 can be used to predict the upcoming NP2 in verb-medial (NVN) (Exp.1) and verb-final spoken sentences (NVN) (Exp.2).

We used a visual-world task modelled on prior studies with adults (Kamide, Scheepers & Altmann, 2003). The critical scenes depicted three objects: the object labelled by NP1 (e.g., rabbit), a potential Theme (e.g, carrot), and a potential Agent (e.g., fox). The visual scene was accompanied by a spoken utterance, where the first argument was either in nominative or accusative case, as in (1) and (2). There were 25 children (4;6-5;0) and 39 adults for Exp.1; 31 children (4;6-5;0) and 21 adults for Exp.2.

(1) Exp.1, NVN (SVO vs. OVS):
Tavşan-   / -ı   birazdan   yiy-ecek   şurada-ki   havuc-u   / tilki.
rabbit-Nom /-Acc   shortly   eat-Fut   there-Rel   carrot-Acc / fox-Nom
(2) Exp.2, NNV (SOV vs. OSV):
Tavşan-   / -ı   birazdan   şurada-ki   havuc-u   / tilki   yiy-ecek.
rabbit-Nom /-Acc   shortly   there-Rel   rabbit-Acc / fox-Nom   eat-Fut
SVO & SOV: 'The rabbit will shortly eat the carrot over there.'
OVS & OSV: 'The fox over there will shortly eat the carrot.'

We analyzed a time window corresponding to the modifier preceding the second noun (underlined). Our dependent variable was Agent Preference (% time looking to Agent – % time to Patient). In Experiment 1, both children and adults had a greater Agent preference in the accusative condition ($p < .05$) indicating that they were able to use case to predict the upcoming noun when verb information was also available. But critically, this effect persisted in Experiment 2 when the modifier preceded the verb ($p < .003$).

Thus, Turkish-speaking children (and adults) can use case markers in the absence of verb information to predict upcoming arguments. This is consistent with a theory of language comprehension and development, which posits that children apply adult-like incremental parsing routines and rely on cues that are both frequent and valid (Trueswell & Gleitman, 2007; MacWhinney & Bates, 1989). Such a theory also provides a satisfying explanation for the failure of German children to integrate case during comprehension.

Sarcasm: Do you hear it now?

Sara A. Peters[1,2], Kathryn Wilson[1] & Amit Almor[1] ([1]University of South Carolina, [2]Newberry College)
sara.ann.peters@gmail.com

Native English speakers (*N.S.*) identify sarcasm in spoken language using a combination of prosodic cues and context. However, it is unclear if non-native speakers (*N.N.S.*) of English use the same information or rely on other strategies. The current work aimed to determine whether *N.N.S.* can identify sarcasm as well as *N.S.*, and also to determine whether there are any differences in the weighting of context and prosody.

We created 30 experimental items consisting of 3-sentence spoken discourses, intermixed with 30 fillers. The first sentence introduced a character and action (see Table), the second sentence introduced a context (either Positive or Negative), and the third sentence presented a second character's feeling about the action (stated using either Sarcastic or Sincere prosodic cues). Sarcastic prosody was generated based on cues identified as important to *N.S.* (slower tempo, greater intensity, lower pitch; Rockwell, 2000). Participants answered a comprehension question, indicating what they thought the second character thought of the first's actions. Answers to the questions were dependent on participants' interpretation of the items and were expected to vary by condition. For the Positive-Sincere condition, interpretations should have been sincere (Yes), and for the Negative-Sarcastic condition, insincere (No). For Positive-Sarcastic, or Negative-Sincere, answers were hypothesized to vary depending on whether listeners judged context or prosody cues as being more important. We hypothesized that *N.N.S.*, following input-processing models of language acquisition, would always value context over prosody.

We recruited 25 *N.S.* and 45 intermediate *N.N.S.* Our analysis consisted of mixed-effects models where response to the comprehension question (Yes/No) was the outcome variable. We found that *N.S.* were better overall than *N.N.S.* at identifying both sincerity and sarcasm[1]. Also, *N.S.* tended to weigh the context more heavily than *N.N.S.* when the prosody did not agree with it[2]. Lastly, *N.N.S.* only used prosody information when the context was Negative, while in Positive contexts it was disregarded[3]. Additional model testing indicated that these findings could not be explained by native language, age, years of English instruction, age of first exposure to English, or time in an English language environment.

Overall, this work suggests that *N.N.S.* rely on different cues than *N.S.* to detect sarcasm in English. Specifically, this work supports an input processing view of language acquisition (VanPatten, 2008), which predicts that learners will focus on lower-level input like context (supplied by lexical information) before higher-level cues like syntax and prosody.

	Positive Context	Negative Context
Sincere Prosody	John offered to go to the store. He picked up the bread, and also remembered the milk. Angie thanked John for doing such a great job.	John offered to go to the store. He picked up the bread, but forgot the milk. Angie thanked John for doing such a great job.
Sarcastic Prosody	John offered to go to the store. He picked up the bread, and also remembered the milk. Angie thanked John for doing ***such a great*** job.	John offered to go to the store. He picked up the bread, but forgot the milk. Angie thanked John for doing ***such a great*** job.
?	Did Angie think John did a good job? (Yes) or (No) Yes answers to comprehension questions always indicated perceived sincerity, while No answers indicated a sarcastic interpretation of the second speaker's judgment.	

[1]Sincerity - 96.5% vs. 75.8% "Yes" responses, $\beta = 0.207$, $p < .001$; Sarcasm - 82.0% vs. 38.1% "No" responses, $\beta = -0.647$, $p < .001$. [2]$\beta = -0.102$, $p < .05$, $\beta = -0.501$, $p < .001$. [3]Only after a Negative Context did Prosody have a significant effect for N.N.S. ($\beta = -0.115$, $p < .001$, $\beta = -0.139$, $p < .001$).

References [1] Rockwell, P. (2000). Lower, slower, louder: Vocal cues of sarcasm. Journal of Psycholinguistic Research, 29(5), 483-495. [2] VanPatten, B. (2008). Input processing in adult second language acquisition. In B. VanPatten & J. Williams (eds.), Theories in second language acquisition: An introduction, 115-135. New York: Rutledge.

Poster Session 2

L2 processing of Arabic derivational morphology

Suzanne Freynik & Polly O'Rourke (University of Maryland)
freynik@umd.edu

Research has shown that L2 learners are less sensitive than native speakers to morphological structure [3,4]. The majority of the research on L2 morphological processing has focused on inflectional morphology. L2 processing of derivational morphology has received less attention and most of the research on it has been limited to Indo-European languages in which derivational morphology is much less productive than inflectional morphology [5]. In contrast, Semitic languages like Arabic and Hebrew have rich, complex and productive systems of derivational morphology involving a root and a template structure which are interleaved to form words. Studies show that native speakers of Arabic process its derivational morphology in a manner similar to that of native speakers of other languages studied [1,2] in that they decompose morphologically complex words during lexical processing.

The current study brings evidence of morphological priming in proficient L2 learners of Arabic (N = 28) to bear on the controversy surrounding the source(s) of the observed differences between L1 and L2 morphological behavior. A lexical decision task with cross-modal priming was used to test whether L2 learners of Arabic engage in morphological processing during lexical access. Morphological, phonological and semantic relationships between the prime and target words were manipulated to test for priming in prime-target pairs sharing the same root morpheme. We analyzed response times using ANOVAs with language group as a between-subjects factor, and priming conditions and all interactions as within-subjects factors. The effects of morphological root priming were shown to be significant (p<0.01), such that reaction times were faster when the prime and target shared the same root (mean 1069.9ms) than when they did not (mean 1182.7ms). This effect did not interact with language group, semantic relatedness or phonological similarity. The results demonstrate that L2 learners of Arabic are able to process and make use of Arabic morphology during lexical access, despite the numerous typological differences between Arabic and their L1 (English), and even in the absence of semantic relatedness. We interpret these findings as evidence against theoretical approaches like the Shallow Structures Hypothesis [3] that attribute observed differences in L2 morphological processing to L2 learners' inability to represent and/or access sublexical representations.

References

[1] Boudelaa, S., & Marslen-Wilson, W. D. (2000). Non-concatenative morphemes in language processing: Evidence from Modern Standard Arabic. In *ISCA Tutorial and Research Workshop (ITRW) on Spoken Word Access Processes*.
[2] Boudelaa, D. S., & Marslen-Wilson, W. D. (2005). Discontinuous morphology in time: Incremental masked priming in Arabic. *Language and Cognitive Processes*, *20*(1-2), 207-260.
[3] Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*(1), 3.
[4] Neubauer, K., & Clahsen, H. (2009). Decomposition of inflected words in a second language. *Studies in Second Language Acquisition*, *31*(03), 403-435.
[5] Silva, R., & Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, *11*(2), 245.

Referential ambiguity and pronoun resolution: Evidence from pupillometry
Manizeh Khan & Jesse Snedeker (Harvard University)
khan@wjh.harvard.edu

Adults readily interpret pronouns, applying strong biases to determine reference in ambiguous cases (Garnham, 2001, for review). For example, in many discourse contexts (Kehler, Kertz, Rohde, & Elman, 2008), subject pronouns are preferentially interpreted as referring to the subject of the prior sentence (e.g. Gernsbacher & Hargreaves, 1988; Järvikivi, van Gompel, Hyönä, & Bertram, 2005). Evidence from ERP paradigms has shown that referential ambiguity in the local discourse context can lead to a processing cost for pronouns (Nieuwland & van Berkum, 2006). The current study builds on prior work by comparing pronoun processing across different task contexts. Is there always a processing cost to ambiguity or do strategies vary depending on global properties of the context, such as whether earlier sentences have violated discourse biases? Pupil dilation is used as a measure of processing cost in two experiments that vary in terms of whether ambiguous pronouns are disambiguated later in the sentence.

Experiment 1. 32 participants listened to 16 two-sentence vignettes, such as: *Diego went to the park with Arthur last Saturday. He ran around playing with Arthur/Diego's dog and then went over to the slide and the swings.* At the possessive, *Arthur's* or *Diego's*, the previously ambiguous pronoun can be resolved to the other character, allowing the subject bias to be violated for half of the ambiguous pronouns. Ambiguity and pronoun resolution (1st versus 2nd mentioned), were manipulated in a 2x2 design. Using a cluster-based permutation test (Maris & Oostenveld, 2007), we uncovered time-regions with significant effects of ambiguity and pronoun resolution on pupil dilation. Ambiguous pronouns led to significantly greater pupil dilation than unambiguous pronouns (Figure 1, $p=.03$, 1500-2900ms after pronoun onset). A significant interaction of ambiguity and pronoun resolution emerged 1300-3000ms after disambiguation ($p=.02$) because pupil size decreased if a subject bias was confirmed and continued to increase if disambiguation indicated a 2nd mentioned resolution.

Experiment 2. The disambiguating word was removed from the vignettes (*a friendly dog* instead of *Diego's dog*. This allows us to see if the ambiguity cost observed in Experiment 1 was due to the knowledge that bias-violating material might be presented later in the sentence. A gender-violation condition was also included, e.g. *she* in a story about boys. As expected, violation pronouns led to significantly greater pupil dilation than ambiguous pronouns (Figure 2, $p<.01$, 800-3000ms following pronoun onset). However, there was no effect of ambiguity ($ps>.5$). Further, there was a significant interaction between ambiguity and Experiment ($p=.01$). Critically, in their offline responses, participants overwhelmingly preferred subject resolutions for the ambiguous pronouns (M=89.8%), demonstrating that they still made referential commitments despite the lack of disambiguation in the discourse.

Together, these experiments demonstrate how global context affects pronoun processing. In a context where biases could be violated downstream, there was a processing cost for ambiguity. When disambiguation was not provided, there was no such cost. One explanation is that the absence of disambiguation resulted in shallow processing, such that referential commitments were delayed (Stewart, Holler & Kidd, 2007). However, the current study involved offline comprehension questions that probed pronoun resolution, which should discourage shallow processing. Another explanation is that listeners might automatically assign subject reference to ambiguous pronouns, and never realize the ambiguity, if their biases are never violated.
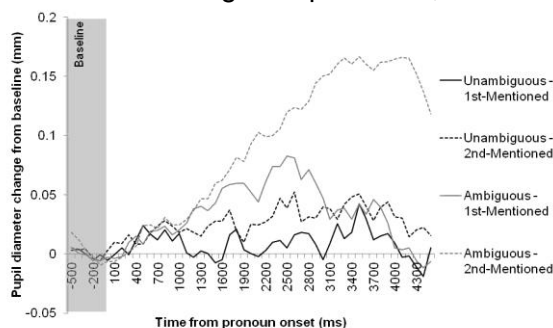

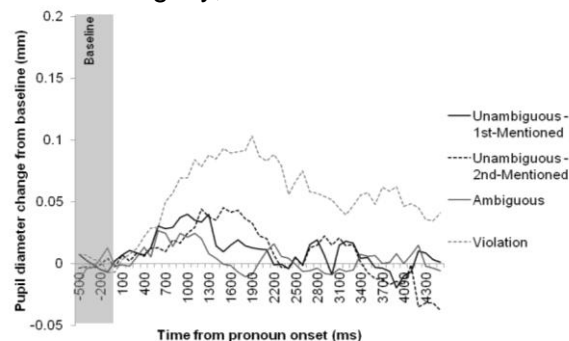*Figure 1.* Pupil response to unambiguous and temporarily ambiguous pronouns.

Figure 2. Pupil response to unambiguous, globally ambiguous, and gender-violation pronouns.

The effect of phrase length on the form of referring expressions

Hossein Karimi[1], Kumiko Fukumura[2], Martin Pickering[3] & Fernanda Ferreira[1] ([1]University of South Carolina; [2]University of Strathclyde; [3]University of Edinburgh)
karimi@email.sc.edu

When people refer to a previously-mentioned Noun Phrase (NP) in discourse, they can choose between different forms of referring expressions such as a repeated noun ("the actor") or a pronoun ("he"). The driving force behind this choice has been argued to be conceptual accessibility, with more accessible NPs being referred to by less marked referring expressions (e.g., a pronoun rather than a repeated noun). Phrase length has been shown to influence conceptual accessibility but its potential effect on the form of referring expressions has not been investigated. Two competing predictions can be made about how length on the part of the referent (referent length) may affect the referring expression employed to realize it: 1. Shorter NPs may be interpreted as more 'given' and therefore may be rendered more accessible, 2. Long NPs may be rendered more accessible because of the greater amount of information predicated on them. Likewise, two predictions can be made about how length on the part of the competitor (i.e., competitor length) may affect the form of referring expression used to realize the referent: 1. Competitor length may reduce referent activation because it might be costly for the processor to keep multiple NPs activated simultaneously, 2. Competitor length may not have any effect on how the referent is realized because discourse models may not be actively updated.

Two written and an oral sentence completion experiments investigated the effect of length on the choice between pronouns and repeated nouns. In Experiment 1 (written), participants read sentences such as "The actor *who was frustrated and upset about the last night's disastrous performance* walked away from the actress *who was frustrated and upset about the last night's disastrous performance.*" and had to come up with a continuation. The results replicated the well-established effect of NP position (more pronouns for NP1 than NP2) but failed to show a clear effect of length. In Experiment 2 (written), the genders of the two NPs were kept the same to increase the likelihood of repeated noun use. The results showed an interaction between length and NP position, suggesting that competitor length reduces pronoun use for the referent (Figure 1). Experiment 3 (oral) was a direct replication of Experiment 1 except that it was done orally. The motivation was that oral production might impose a greater burden on short-term memory and therefore might affect processing. The results of Experiment 3 showed a referent length effect, with longer NPs being more likely to be realized with a pronoun, but failed to show a competitor length effect (Figure 2). Moreover, across the experiments, a significant modality effect was observed, with significantly more pronoun use in the written (Experiment 1) than in the oral modality (Experiment 3).
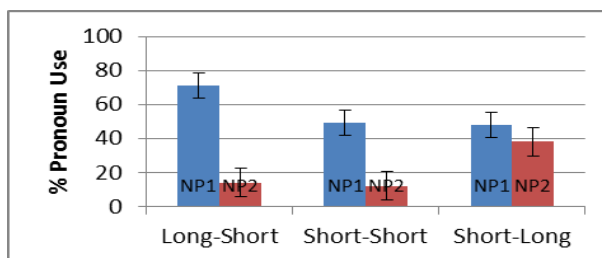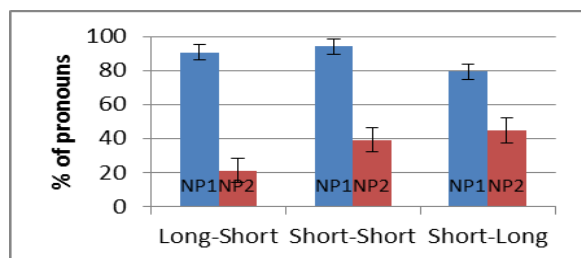

Figure 1. Results for Exp 2 (written production)   **Figure 2.** Results for Exp 3 (oral production)

Taken together, the results seem to suggest that both referent length and competitor length affect form of referring expressions but these effects depend on the specific strategies employed for the specific task at hand.

The myth of the Overt Pronoun Constraint in Spanish

Carlos Gelormini (Institute of Cognitive Neurology[1], David Huepe[2], Eduar Herrera[3], Timothy W. Boiteau[4], Margherita Melloni[1], Facundo Manes[1], Adolfo Garcia[5], Agustin Ibañez[1] ([1]Institute of Cognitive Neurology, Buenos Aires; [2]Universidad Diego Portales; [3]Universidad Autónoma del Caribe; [4]University of South Carolina, [5]Universidad Nacional de Córdoba)
cgelormini@gmail.com

The Overt Pronoun Constraint (OPC) states that, in null subject languages, overt pronouns cannot be bound by *wh-* or quantifier antecedents. In a Spanish question such as *¿quién cree que él va a ganar?* [who thinks that he will win?], the overt pronoun *él* [he] can only be anaphoric: it must refer to someone mentioned earlier in the discourse. The bound reading (*who thinks of himself as a possible winner?*) is not available. Similarly, in *cada estudiante cree que él va a ganar* [each student thinks that he will win] the overt pronoun *él* [he] must be anaphoric. The bound interpretation (*each student thinks of himself as a possible winner*) is not available. The OPC has for decades been used to show that L1 and L2 grammars are UG constrained: if the OPC is not taught nor transferred from the L1, then it must be a universal principle. OPC obedience in L1 has been taken for granted. Through two written questionnaires, we examined the OPC in 246 native speakers of three dialects of Spanish: Colombia, Chile, and Argentina.

In Exp. 1, items consisted of a context and a question with an overt pronoun and a *wh-* antecedent. Participants chose one of two possible answers and rated their choice on a confidence scale from 1 to 10. *Sample Item - Context:* Ninguna actriz admitió haberse equivocado durante la función. Un actor admitió que la directora se equivocó al no estar presente en el ensayo general. [No actress admitted having made a mistake during the performance. An actor admitted that the female director had made a mistake by not being present at the dress rehearsal]. *Question:* ¿Quién admitió que ella cometió una equivocación? [Who admitted that she made a mistake?] *Answer:* (a) un actor [an actor]; (b) nadie [nobody]. Answer (a) implied that the overt pronoun had an anaphoric interpretation whereas answer (b) implied a bound interpretation. In Exp. 2, participants judged sentences with overt pronouns and quantifiers as true or false according to a context and rated their answers on a confidence scale from 1 to 10. *Sample Item - Context:* Ningún jugador del equipo reconoció haber fallado. Un jugador reconoció que el técnico había fallado en su trabajo. [No player on the team admitted making a mistake. One player admitted that the coach had made a mistake]. *Sentence:* Ningún jugador reconoció que él había fallado. [No player admitted that he had made a mistake]. In this case, a bound reading of the overt pronoun made the sentence true whereas an anaphoric reading made the sentence false. Unambiguous filler items were used in both experiments.

Results showed a similar number of bound and anaphoric answers in Exp. 1 and a higher number of bound answers in Exp. 2. In Exp. 1, a main effect of choice ($F(2, 256)=67.56$, $p<0.0001$, $\eta_2= 0.34$) with post hoc comparisons (Tukey HSD, MSE = 1189.0, df = 256.00) showed that correctly classified fillers yielded the highest percentage relative to bound and anaphoric ($ps<0.001$) answers. Importantly, no differences between bound and anaphoric answers were observed. A main effect of confidence rate ($F(2, 196)=11.45$, $p<0.001$; $\eta_2= 0.10$) with post hoc comparisons (Tukey HSD, MSE = 1.6414, df = 196) showed that fillers yielded higher confidence rates than bound ($p<0.05$) and anaphoric ($p<0.001$) answers. In Exp. 2, a main effect of choice ($F(2, 226)=89.37$, $p<0.001$, $\eta_2=0.44$) with post hoc comparisons (Tukey HSD, MSE = 1096, df = 226) showed that anaphoric answers had a lower percentage than bound answers ($p<0.001$) and correctly classified fillers ($p<0.001$). A main effect of confidence rate ($F(2, 172)=4.39$, $p=0.01$, $\eta_2=0.04$) with post hoc comparisons (Tukey HSD, MSE = 1.66, df= 172) confirmed that anaphoric answers yielded lower confidence rates than fillers ($p<0.001$).

To conclude, sentences containing overt pronouns and *wh-* or quantifier antecedents are ambiguous: bound and anaphoric readings are possible. This study challenges the widely held myth of the OPC, which should no longer be used as an argument for the involvement of UG in L1 or L2 grammars. Our study also shows that simplistic configurational formulas fail to make the right predictions about native speakers' interpretation of overt and null pronouns.

What types of lexical information are reaccessed during pronoun processing?

Sol Lago[1], Shayne Sloggett[2], Wing Yee Chow[1] & Colin Phillips[1] ([1]University of Maryland; [2]University of Massachusetts)
sollago@umd.edu

Previous studies have demonstrated lexical frequency effects in the processing of coreference: pronouns with infrequent antecedents are processed more slowly than pronouns with frequent antecedents, suggesting that processing a pronoun involves re-accessing the lexical representation of its antecedent [1-3]. In production, frequency effects in word naming have been linked to activation of the form (*lexeme*) of a word [4]. One might therefore expect that in comprehension, antecedent frequency effects in pronoun processing should be accompanied by reactivation of the antecedent's form. Alternatively, the coupling of frequency with form might be specific to production. In this study, we examined whether semantic information, phonological information or both affect pronoun processing. Our results show that, in contrast with production, frequency effects co-occur with semantic effects, but not phonological effects. This replicates previous findings that have shown semantic reactivation of the antecedent of a pronoun [5-6], and suggests a contrast between comprehension and production with regard to whether frequency effects are uniquely associated to the form of a word.

In the present eye-tracking experiment (n=40) we varied the antecedent of a pronoun such that it was either related or unrelated to the word immediately following the pronoun, where this relationship could either be phonological/orthographic or semantic. We predicted that the retrieved properties of the antecedent should impact reading times when they overlapped with the word following the pronoun. To isolate relatedness effects specifically due to referential processing, we manipulated whether a pronoun or a determiner preceded the target word, resulting in a 2 (related/unrelated) × 2 (semantic/phonological) × 2 (pronoun/determiner) design. To verify that the lexical representation of the antecedent was re-accessed, we regressed the lexical frequency of the antecedent against reading times: frequency was a significant predictor at the antecedent region in early measures (*first fixation:* β=-10.16, $p<.05$; *first pass:* β=-13.275, $p<.09$), and consistently with the lexical re-access account it was also a predictor at the crucial post-anaphor region in late measures (*re-read times:* β=-67.88, $p<.05$; *total times:* β=-68.02, $p<.05$). Furthermore, in the semantic conditions, relatedness led to shorter re-read times at the antecedent (β=-82.36, $p<.05$) and post-anaphor regions (β=-113.33, $p<.01$) only in the pronoun conditions. In the phonological conditions, however, no effect of relatedness was found, and only a main effect of pronoun was observed, with the pronoun conditions more likely to elicit regressions from the post-anaphor region than the determiner conditions (β=.43, $p<.05$). Taken together, these results show that processing a pronoun involves lexical re-access (as evidenced by antecedent frequency effects), but that in contrast with production, antecedent reaccess does not entail activation of phonological/orthographic information, and that lexical frequency information is not uniquely tied to form representations.

Semantic conditions
Related. The maintenance men told the **singer** about a problem. They had broken *the/his* **piano** and would have to repair that first.
Unrelated. The maintenance men told the **deputy** about a problem. They had broken *the/his* **piano** and would have to repair that first.

Phonological conditions
Related. The maintenance men told the **singer** there would be a delay. They said that *the/his* **sink** wouldn't be installed until next month.
Unrelated. The maintenance men told the **deputy** there would be a delay. They said that *the/his* **sink** wouldn't be installed until next month.

References. [1] Finocchiaro & Caramazza (2006) *LCP*; [2] Lago, Chow & Phillips (2011) *CUNY*; [3] Schmitt, Meyer & Levelt (1999) *Cognition*; [4] Jescheniak & Levelt (1994) *JEP: LMC*; [5] Nicol & Swinney (1989) *Journal of Psycholinguistic Research*; [6] Shillcock (1982) *Language and Speech*.

Disfluency Primes
Sarah Brown-Schmidt (University of Illinois)
sarahbrownschmidt@gmail.com

According to one theory, disfluencies are word-like symbols speakers select to communicate specific meanings (Fox Tree & Clark, 1996). Indeed, listeners capitalize on disfluency to make inferences about upcoming referents (Arnold, et al., 2004; Brennan & Schober, 2001). If disfluency is structured and selected, disfluency should be subject to priming effects, much like lexical and syntactic priming. Alternatively, if disfluency is solely a by-product of production difficulty, disfluencies should result from difficulty and should not prime, because whether or not a speaker was primed on a given trial would be orthogonal to whether they were experiencing production difficulties.

To test this hypothesis, we examined the production of size-modified NPs in situations where speakers spontaneously use prenominal disfluency (e.g., *thee uh large dog*), and repairs (e.g., *the dog, uh large one*), to incorporate size information into the message. Such disfluencies typically result from delays in noticing the scene contains a contrast (a smaller dog; Brown-Schmidt & Tanenhaus, 2006). We asked whether priming different types of disfluency would modulate the rate of each disfluency type.

Experiment 1 established a task to elicit disfluency: An eye-tracked speaker (S, n=36) and experimenter (E) took turns describing objects in a complex display. On critical trials, two objects moved slightly on S's display (e.g., a dog and cup), one of which had a size-contrast on the screen (e.g., smaller dog). S described the objects, e.g. "*the big dog and cup are flashing*". <u>Results</u>**:** Ss produced prenominal disfluencies (e.g., *thee uh big dog…*) and size repairs (e.g., *the dog, uh big one*) frequently, affecting 12% and 27% of NPs respectively. Whether S fixated the size-contrast predicted whether S would modify the critical NP (z=17.89), indicating the fixation indexed the when size information was included in the linguistic message. Analyses of the timing of the first contrast fixation indicated that both types of disfluency were associated with significant delays in incorporating size into the message. Relative to fluent, modified NPs, contrast-fixations were delayed by 371ms (t=3.50) for prenominal disfluency, and delayed by 1701ms for repairs (t=13.02), with similar delays at each NP position (e.g., dog/ cup). These findings suggest a flexible use of disfluency in message planning.

Experiment 2 examined whether these disfluencies could be primed: Ss (n=36) performed the identical task as Experiment 1 and were randomly assigned to a between-subjects priming condition: During E's turn to speak, 75% of the time E produced a prenominal disfluency (e.g., *thee uh large dog…*), a repair (e.g., *the dog uh large one…*), or a fluent control (e.g., *the large dog…*). If disfluency primes, comparisons between the two prime conditions will determine if priming is specific to disfluency type, or whether hearing disfluency increases all types of disfluency. <u>Results</u>**:** Repair priming was significant and specific, resulting in a 13% increase in repairs following repair vs. prenominal disfluency primes (z=2.27), regardless of NP position. Prenominal disfluency priming was not significant (z=.47). While the interaction of disfluency type and NP position for prenominal disfluency priming was not significant (z=1.91), the effect was clearly more pronounced at the first of the two NPs (12% increase over repair primes) vs. the second NP (-6% vs. repair primes). The relationship between disfluency and the timing of the first fixation to the contrast was unchanged by priming (ts<1).

Conclusions: These findings present the first evidence that at least one type of disfluency—size repairs—are sensitive to priming. While disfluencies are clearly related to planning difficulty, the fact repairs were primed suggests some types of disfluency are better seen as productive constructions used to adjust message information, rather than simply an unintended error (Clark & Wasow, 1988). After all, if disfluency was solely a by-product of an error process, it should not be influenced by the behavior of one's communicative partner. Instead, the present results are more consistent of a view of disfluencies as word-like signals that are selected by the speaker. The fact repairs involve syntactic change (e.g., -*small one*) may contribute to the success of repair primes, with a mechanism similar to noun-adjective order priming for fluent constructions (Cleland & Pickering, 2003). Importantly, the present results show priming is not restricted to fluent/grammatical forms. Moreover, the fact disfluency primes emphasizes that disfluency is not solely the result of production problems.

Poster Session 2

Contextual effects on the comprehension of speaker corrections: An ERP study

[1]Justine VanDyke-Lyon, [2]E. Matthew Husband, [1]Fernanda Ferreira, & [3]Nathan Maxfield ([1]Univeristy of South Carolina, [2]St. Hugh's College, Oxford, [3]University of South Florida)
vandykej@mailbox.sc.edu

Current models of language processing assume interpretations are built incrementally. But what happens when our initial representation turns out to be incorrect and an alternative interpretation must be constructed? Research investigating garden-path effects has revealed that the ease with which a structure can be revised is influenced by the strength of the initial representation. Interpretations that are semantically coherent, for example, are harder to relinquish (Ferreira & Clifton 1986). Here we explore whether similar mechanisms operate when speakers must revise an initial interpretation following a speaker self-correction (*The little boy went to the zoo where he saw a lion I mean a tiger*). Repair processes following a correction involve at least two stages that may be affected by the strength of the initial representation: (1) access of the repair, and (2) replacing the word spoken in error (reparandum) with the repair. Using event-related potentials (ERPs), we investigate the extent to which lexico-semantic factors related to the context and the relationship between the words in the reparandum and the repair affect these processes.

The experimental conditions are shown in Table 1. Context was manipulated such that reparanda were embedded in either strongly (e.g., 1 & 2) or weakly (e.g., 3 & 4) constraining contexts. We also manipulated the strength of the relationship between the reparandum and the repairs, which were either highly (e.g., 1 & 3) or weakly related (e.g., 2 & 4). Both manipulations were assessed using Latent Semantic Analysis. If these factors affect repair processes, we expect to see amplitude differences in ERP components that are associated with the repair stages outlined above; thus, analysis focused on ERP components whose temporal and spatial characteristics were consistent with: (1) N400, reflecting lexical access and (2) late positivities, reflecting revision/integration processes, as established in fluent language paradigms (Federmeier et al, 2007).
We recorded ERPs from 26 adults. Participants heard 128 randomly presented sentences comprising 4 conditions. ERPs were time-locked to the onset of 3 critical events: the reparandum, the edit term, and the repair. The primary focus for analysis was the repair. Results from a traditional windowed amplitude analysis as well as a temporal-spatial principal components analysis revealed: (1) a smaller central negativity peaking at 428ms to highly-related compared to weakly-related repairs, with the effect most pronounced in strongly constraining contexts; (2) a larger central positivity peaking at 674ms to highly related repairs in weakly compared to strongly constraining contexts.

These results suggest that the local, semantic relationship between reparanda and repairs affects the ease with repairs are accessed. Subsequent revision/integration, however, is further influenced by the extent to which concepts have been activated by context. Although highly-related repairs appear to be more effortful to integrate, consistent with our predictions, this process is made even harder if context has not provided prior activation for a concept, contrary to our predictions. Discussion focuses on the functional significance of the late-positivity observed as a reflection of global conceptual activation and integration (Brouwer et al., 2012). We conclude that the processing of disfluencies is contingent on the same mechanisms that are involved in the processing of fluently-spoken language.

Condition	Context	Reparanda	Edit	Repair
(1) Strongly constraining/ highly related	The little boy went to the zoo where he saw	a lion	I mean	a tiger
(2) Strongly constraining/ weakly related	The little boy went to the zoo where he saw	a cougar	I mean	a tiger
(3) Weakly constraining/ highly related	The little boy went on a trip where he saw	a lion	I mean	a tiger
(4) Weakly constraining/ weakly related	The little boy went on a trip	a cougar	I mean	a tiger

Lexical disambiguation using parafoveal information

Rukshin Shaher and Shravan Vasishth (University of Potsdam)
rukshin@gmail.com

Several studies have shown parafoveal-on-foveal effects in reading which provide some evidence in support of the parallel attention shift or attention gradient models (Richter et al., 2006) of eye movement control. However there is conflicting evidence for whether semantic information can be extracted from the parafovea, and whether such information can affect foveal processing.

In the current experiment, we examine the effect of semantic parafoveal preview on the processing of ambiguous nouns in English using the boundary paradigm. We used sentences such as in (1) and manipulated (i) the nature of the parafoveal cue (in 1a and 1b, the cue 'electrician' suggests the preferred reading of the ambiguous word 'fans', and in 1c and 1d, the cue 'player' suggests the dispreferred reading of the ambiguous word), (ii) the resolution of the ambiguous word in the sentence-final region (preferred reading in 1a and 1d, and dispreferred reading in 1b and 1c), and (iii) whether a preview of the parafoveal word was provided or not.

34 participants were asked to read 102 sentences (32 items and 70 fillers). One word in each of these 102 sentences was replaced by x's. The position of the x'ed out word was varied for the filler sentences. For the no preview conditions, the semantic cue was x'ed out when the sentence appeared on the screen while for the preview conditions, the cue was x'ed out only after the eyes moved out of the ambiguous region.

The ambiguous words is in **bold face**, the parafoveal cue is *italicized*.

1.   **(1) a.   Preferred meaning cue + Preferred Resolution**
2.   Despite the loud noise of the **fans**, the *electrician* insisted that they were functioning correctly.
3.   **b.   Preferred meaning cue + Dispreferred Resolution**
4.   Despite the loud noise of the **fans**, the *electrician* managed to communicate with the player over their cheering.
5.   **c.   Dispreferred meaning cue + Dispreferred Resolution**
6.   Despite the loud noise of the **fans**, the *player* managed to communicate with the coach over their cheering.
7.   **d. Dispreferred meaning cue + Preferred Resolution**
8.   Despite the loud noise of the **fans**, the *player* preferred to keep them switched on rather than suffer the heat.

We found an interaction (t=2.46) of preview and parafoveal cue on single fixation durations in the ambiguous region: In the preview conditions, the ambiguous region was read faster when the parafovea provided a preferred cue. In addition, in the disambiguating region at the end of the sentence, there was (i) a main effect (t=2.83) of whether the parafoveal cue was previewed or not, (ii) a main effect (t=2.99) of the resolution preference of the ambiguous word, and (iii) an interaction (t=2.77) of the nature of the cue and the resolution preference of the ambiguous word.

These results suggest that semantic information can be extracted from the parafovea and that this information can have an effect on the processing of the foveal word. These results are in line with the parallel attention shift models of eye movement control and go against the predictions of the serial attention shift models (Reichl et al., 2009).

References

Reichl, E. D., Warren, T., McConnell, K. (2009). Using EZ reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review, 16(1),* 1-21.
Richter, E. M., Engbert, R., Kliegl, R. (2006). Current advances in SWIFT. *Cognitive Systems Research, 7(1)*, 23-33.

Morphological activation during spoken word recognition in Hebrew

Daphna Heller (University of Toronto) & Avital Deutsch (Hebrew University of Jerusalem)
daphna.heller@utoronto.ca

In Indo-European languages derivational morphology involves concatenation of morphemes (e.g. dark-ness). In Semitic languages, like Hebrew and Arabic, (most) words are constructed by *interweaving* two morphemes. For example, *mexlaf* "interchange" is composed by inserting the consonants of the root {x,l,f} into the vocalic pattern meXXeX, (X are the position of the consonants). Because roots in Semitic do not have an independent phonological form, their status during spoken word recognition is of particular interest.

Evidence about the activation of roots during lexical access comes from priming experiments with visual words, where facilitation of lexical decision and naming were observed when the prime shared the same root (e.g. Frost et al, 1997 for Hebrew; Boudala & Marslen-Wilson 2000 for Arabic). Importantly, this facilitation was relative to a condition where the prime overlapped with the target in three non-root letters. In addition, this effect was found even when the semantic relationship between the prime and the target was opaque. However it is not immediately obvious whether this will extend to the auditory modality, because it is possible that the consonantal nature of the writing systems in Semitic languages (where most vowels are not marked) gives consonants a special status above and beyond their auditory prominence.

Here we use the visual world eye-tracking paradigm to explore the activation of roots during spoken word recognition in Hebrew. Participants saw an array of four pictures and (after a delay of 800ms) heard a single word – participant had to click on the corresponding picture. Critical trials contained the target picture, a competitor picture, and two unrelated pictures. In the morphological condition, the competitor shared the root with the target word, e.g. target: me**xl**af "interchange", competitor: **x**ali**f**a "suit" (the unrelated words did not share any consonants with the target). In the phonological condition, the competitor shared three consecutive consonants with the target, but these consonants were not the root of the target, e.g. target: sa**rgel** "ruler", competitor: **regel** "leg" (the Hebrew lexicon forced a between-item design). In addition, target and competitor pairs were not cohorts or rhymes, and their semantic relationship was opaque.

We examine the likelihood of making saccades to the different objects; data was analyzed using mixed-effects logistic regression models with random intercepts for subjects and items (we report data from 16 native Hebrew speakers). First, we examine the likelihood of making a saccade to the competitor vs. an unrelated distractor in the two conditions during the processing of the target word (mean length 710ms). In the morphological condition, participants were more likely to launch a saccade to the root-related competitor compared with an unrelated distractor (β=-0.34, *SE*=0.15, *z*=-2.33, *p*=0.02), but in the phonological condition, the competitor did not receive more saccades than an unrelated distractor (β=-0.20, *SE*=0.15, *z*=-1.34, *p*=0.19) These results indicate that morphological activation exceeds what is expected from phonological overlap alone (the short preview makes it unlikely that these are semantic effects – see Yee et al. 2011). Furthermore, examining the time course of this effect reveals that it appears about 500ms after word onset. Given the 200ms estimate for launching a saccade, this means that the elevated saccades to the morphological competitor occurs after hearing 300ms of the word, which is after the second root consonant (average time: 242ms).

These results extend previous findings about root activation during lexical access in Hebrew by showing that this activation occurs during *spoken* word recognition. To our knowledge, this is the first study to show that mere morphological overlap can draw visual attention.

References. [1] Boudelaa, S., & Marslen-Wilson, W.D. (2000). Non-concatenative morphemes in language processing: Evidence from Modern Standard Arabic. [2] Frost, R., Forster, K.I., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked priming investigation of morphological representation. *JEP: LMC* 23, 829–856 [3] Yee, E., Huffstetler, S., & Thompson-Schill, S.L. (2011). Function follows form: Activation of shape and function features during object identification. *JEP General*, 140(3), 348-363.

Predictability and prediction: Are upcoming words pre-activated during sentence processing?

Wonil Choi, Peter C. Gordon (University of North Carolina at Chapel Hill)
wichoi@email.unc.edu

A word in a sentence is processed more easily if it is predictable based on the preceding context than if it is not. This well-established phenomenon can be explained in two very different ways. On one account, predictability leads to predictive processes in advance of the word's appearance. Predictive models have been very prominent in recent theories of human sentence processing (e.g., DeLong, Urbach, & Kutas, *Nature Neuroscience*, 2005). On the second account, the predictable word is not actually predicted, but is easier to process when it occurs because its meaning can be readily integrated with that of the preceding context; on this view, measures of predictability, such as cloze data, are a proxy for plausibility of the meaning of the word combined with the preceding context. Evidence for integrative effects of this sort is found in demonstrations of processing ease varying with the semantic fit of a word even in cases where words are not predictable from preceding context. At issue is whether additional predictive mechanisms are required to account for the processing of predictable words.

Prediction can be distinguished from integration in that the benefit to processing that is found when prediction is accurate should be accompanied by a cost when prediction is inaccurate, whereas the effects of integration depend only on plausibility. Further, prediction should have early effects on processing while integration has been found to have later effects (e.g., Rayner, Warren, Juhasz, & Liversedge, *JEP:LMC*, 2004). The dynamics of the benefit and cost of word predictability were assessed in an eye-tracking study where target and control words, matched in length and frequency (e.g., *pulse* and *watch*), appeared following a nonconstraining context such as 1 where no word was highly predicted (average cloze of most common continuation = 2%) and following a constraining context such as 2 where the target word was highly predictable (average Cloze = 89%). The last two words of the initial sentence context were identical in the nonconstraining and constraining contexts and the meaning of the control word was equally plausible in the nonconstraining and constraining sentence contexts.

1 Jack always remembers to check his _____
2 She held the unconscious man's wrist to check his _____

Three important results are illustrated in the figure to the right: (1) Processing of the target word showed a benefit (shorter gaze durations) in the constraining context (where it was predictable) as compared to the nonconstraining context where it was not. (2) There was no difference in gaze duration between the target and control words in the nonconstraining context, indicating that they were well matched. (3) Processing of the control word showed a cost (longer gaze durations) in the constraining context (where the target was predictable) as compared to the nonconstraining context where no

Gaze Duration on Critical Word

word was highly predictable. The highly significant interaction between the target-control and context constraint factors seen for gaze duration was not found in measures of downstream processing (first-pass reading on subsequent words or measures of later processing on any region). The absence of later effects provides evidence against an explanation in terms of an integrative mechanism as such effects typically occur with some delay. The finding that the benefit of predictability came with a cost is consistent with the operation of a language processing system that makes predictions.
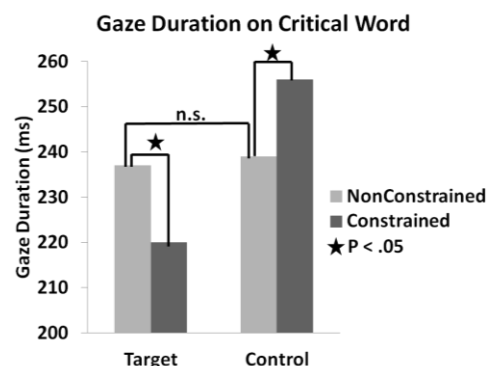
The influence of context information on vocabulary acquisition in reading
Randy Lowell & Robin K. Morris (University of South Carolina)
lowell@mailbox.sc.edu

Unknown words often appear during the course of silent reading and there is evidence to suggest that readers infer meaning from the surrounding context (e.g., Chaffin, Morris, & Seely, 2001; Williams & Morris, 2004). In fact, context information has been shown to enhance acquisition of meaning for unknown words beyond presentation of a dictionary definition alone (e.g., Bolger, Balass, Landen, & Perfetti, 2008; Fischer, 1994). The educational literature has focused primarily on the construction of instructional contexts to assess retention (e.g., McKeown, 1985; Nagy, Anderson, & Herman, 1987; Nagy & Herman, 1984; Nagy & Herman, 1987; Nagy et al., 1985). Less has been done to evaluate aspects of the context that make it more or less informative during implicit learning situations. Further, it is critical to better understand the moment to moment processes that take place during the course of natural reading, because those processes could reflect the development of novel word meaning. In the current study, adult skilled readers encountered informative context followed by a known or novel target word while their eye movements were monitored. Two potential, but not necessarily mutually exclusive, aspects of context informativeness were investigated: contextual constraint (Experiment 1) and action content (Experiment 2). Following the reading session, each participant was invited to complete a surprise posttest regarding novel words and their meaning.

In Experiment 1, the influence of contextual constraint on vocabulary acquisition in reading was examined. It has been suggested that stronger constraint yields stronger novel word acquisition (Borovsky, Kutas, & Elman, 2010); whereas others might claim that moderate constraint yields superior acquisition (e.g., Lampinen & Faries, 1994). By analyzing posttest performance and eye movement behavior during natural reading, those claims were more thoroughly examined here. The informative context in this case was either strong or moderate constraint, followed by a common-known, uncommon-known, or novel target. In general, readers reread the context and target regions more often and/or for longer when the constraint was moderate than when it was strong. A notable exception was that readers spent more time rereading the strong constraint context region after having encountered a novel word. Readers also spent less time on the end of the sentence following strong constraint and a common-known target, but more time following encounters with a novel target. Despite differences in eye movement behavior, strong and moderate constraint yielded comparable success on the surprise posttest. Posttest data was mapped onto corresponding eye movement trials, which revealed that posttest success was predicted by inflated context/target rereading, but more so for moderate constraint.

In Experiment 2, the influence of action content on vocabulary acquisition in reading was examined. It has been suggested in the literature that language comprehension is enhanced by action-related language (e.g., Cacciari et al., 2011; Fischer & Zwaan, 2008). The goal here was to determine if, while holding contextual constraint constant, the level of action content within the informative context impacts the acquisition process and/or product. The second sentence of pairs contained informative context (low vs. high action content) followed by a known or novel target word. Readers spent more first-pass time on novel targets preceded by low action content, but they more often regressed from the end of the sentence toward novel targets when the item contained high action content. Posttest performance was comparably successful across context conditions; however, particularly inflated rereading of background information in the first sentence predicted posttest success for items containing high action content.

Both experiments provide evidence that readers infer novel word meaning from a single exposure in context, in this case pre-target informative context. Critically, manipulation of the context (contextual constraint and action content, respectively) revealed different reading patterns during acquisition that resulted in comparably successful retention.

Lexical clustering in efficient language design

Kyle Mahowald[1], Steven T. Piantadosi[2], Edward Gibson[1] ([1]MIT, [2]University of Rochester)
kylemaho@mit.edu

Much recent evidence suggests that the structural properties of the lexicon are communicatively efficient relative to the capacities and limitations of language processing mechanisms (Flemming 2004; Piantadosi, Tily & Gibson 2011; Graff 2012). Here, we investigate the phonetic relationships between words in the lexicon and compare their relationship to phonotactically-plausible "null" lexica with no communicative or cognitive optimization. If the lexicon is sparser than expected by chance (fewer minimal pairs, greater distance between words, etc.), that would suggest communicative pressure for the lexicon to avoid confusable words. Alternatively, if the lexicon is less sparse than expected by chance, that would suggest other pressures, such as a learning bias towards familiar phonetic patterns or easier processing from the presence of semantically related words that share sound structures.

Experiment 1. The probability of the lexicon containing a word w is dependent on the phonotactic probability of word w in language L. We asked whether that probability also depends on the phonetic properties of all the other words in the lexicon by creating 100 null English lexica, first by generating candidate words from a 3-phone model of English (trained on a monomorphemic subset of pronunciations from Hayes, 2012) and then sampling to match the real monomorphemic lexicon for distribution of word length and phonotactic probability as measured by Hayes' Blick phontactic probability model (Hayes 2012). Thus, the resulting null lexica have the same lengths and phonotactic probability as the real lexicon but differ crucially in that the words are generated independently (given the model) and are thus not optimized relative to one another. We then computed various lexical statistics to see if the real lexicon could have been plausibly generated by the null lexicon model. We can reject this null hypothesis. The real lexicon is clumpier than the null lexica by almost every measure: it has more average minimal pairs per word (real=2.72, sim mean=1.33, z=47.03, $p < .001$), more neighbors per word (real=3.00, sim mean=1.49, z=47.29, $p < .001$), a higher average number of phonemes from the start of the word until the word is unique (real=3.91, sim mean=3.80, z=8.24, $p < .001$), a smaller Levenshtein distance between randomly sampled 4-letter words (real=3.72, sim mean=3.77, z=3.07, $p < .01$), and less entropy over phones (real=11.30, sim mean=11.5, z=20.62, $p < .001$).

Experiment 2. To investigate whether the clumpiness of the real lexicon is plausibly caused by a tendency for semantically related words to be phonetically similar, we tested whether words are phonetically more similar to their antonyms (which differ in meaning but can typically be used in the same syntactic and semantic context) than to other words of the same part of speech. If semantically related words (such as antonyms) are more similar to each other than to other words, that would suggest a plausible driving mechanism for the effect in Experiment 1. For a set of monomorphemic antonym noun, adjective, and verb pairs from Wordnet, we selected one word from each pair and calculated its Levenshtein distance (using CELEX pronunciations) to its antonym and to every other word in the data set of the same part of speech. The average Levenshtein distance between a word and its antonym was 3.64 compared to 3.98 between a word and a non-antonym. This effect of being an antonym (β=-.08, t=-4.11, $p < .001$) was significant by a mixed effect linear model predicting log Levenshtein distance while controlling for length effects and with random intercepts for the words being compared. This property of antonyms suggests that it is plausible that the clumpiness of the lexicon is caused by phonetic clustering among semantically related words.

Conclusion. We argue that these results arise from several interesting properties of language learning and processing mechanisms: it may be easier to learn new words that are phonetically similar to known words (Storkel 2006), phonetic similarity may be used to communicatively to convey rough semantic classes (as in sound symbolism), and language production and/or comprehension mechanisms may prefer re-use of articulatory sequences, over and above probabilistic phonotactics.

A new account of spillover effects in reading: Evidence from parafoveal masking
Michael Shvartsman, Richard Lewis & Satinder Singh (University of Michigan)
mshvarts@umich.edu

Overview. Words that follow high-frequency words are read more quickly than those that follow low-frequency words (Inhoff and Rayner 1986). The E-Z Reader model of eye movement control (Reichle, et al. 2012) explains the effect by appeal to parafoveal preview: higher frequency words are processed more quickly, allowing more time to preview the next word thus reducing the time spent on that word once the eyes move there. This explanation maintains a tight eye-mind link: only words in the fovea or parafovea are processed. We consider an alternative explanation that breaks this tight link and permits continued processing of previous words without direct perceptual input. We provide experimental evidence for this explanation by showing that spillover persists to the same degree when parafoveal preview is eliminated. We demonstrate through modeling that the only parameterization of E-Z Reader that can account for the results breaks the tight eye-mind link and implements the alternative explanation.

Experiment. In the list lexical decision task (LLDT), participants see a list of character strings and make a single response as to whether they are all words. In the gaze-contingent parafoveally-masked variant we present here, all words other than the word currently being fixated are hidden behind a mask (####), denying participants parafoveal preview. 37 students participated in the experiment for course credit, each of which saw 200 trials of the masked LLDT. Half of trials in each block contained six English words; the remainder contained five words and one nonword in a randomly drawn position. All strings were four characters long. Words were drawn from a subset of the Brown Corpus (Kuçera et al. 1967) with bimodally-distributed frequency (centers at 239.2 and 5.6 occurrences per million). Data was compared to data from a previous experiment identical except for the absence of masking.

Results. Linear mixed effects models were fit with intercepts for subject, trial, word, and ROI, slopes for trial and ROI, and a fixed covariate of word-N frequency; significance was tested by model comparison. Significant spillover effects were found in both masked and unmasked data over a variety of standard measures (e.g. first and single fixation durations, total fixation times; see figure below for single fixation duration on word N as a function of N-1 frequency).

Modeling. We used a simplified implementation of E-Z Reader and confirm that it provides poor fits to the masked data even when refit within parameter bounds previously used in the literature (see figure, left). To fit E-Z Reader to the masked data, it was necessary to allow the model to spend substantially more time processing a word after the decision to move the eyes is made (see figure, bottom right). (In E-Z Reader parameter terms, this means letting Δ range well above 1). No other change to parameter bounds is sufficient to recover the spillover effect. The new model exhibits spillover effects in both the masked and unmasked paradigms in a completely different way than the original E-Z Reader: rather than high-frequency words speeding processing on the next word through preview, low-frequency words slow down processing on the next word, even after the eye moves away from them. This break between perception and cognition has implications for models of eye movement control more broadly, and indicates that spillover is not due only to preview or much higher level linguistic processing.
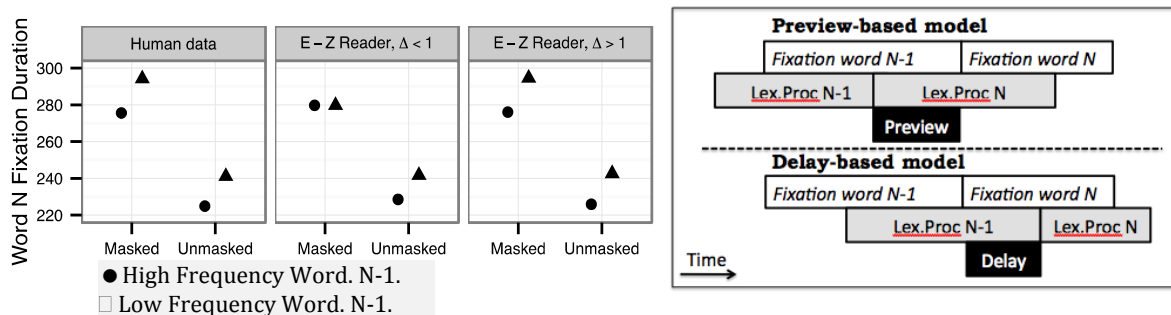
Auditory confusability vs. phonological neighborhood in language production
Susanne Gahl (UC Berkeley), Julia Strand (Carleton College)
gahl@berkeley.edu

It is a widely shared assumption that hearing or saying a target word causes joint activation of non-target words in the lexicon. Yet, the consequences of joint activation remain controversial. For example, while high phonological neighborhood density (PND; the number of words differing from a target by one phoneme, often weighted by the relative frequency of target and neighbors) impedes word recognition (Luce & Pisoni, 1998), the effects of PND on production are unclear: Some studies (e.g. Wright, 2004), based on single-word naming tasks, find high PND to be associated with hyperarticulation (greater vowel dispersion, longer duration) and attribute that pattern to talkers' maximizing the intelligibility of difficult recognition targets. Others (Gahl et al., 2012), based on conversational speech (Pitt et al., 2007), report the opposite (shortening and reduced vowel dispersion for high-PND words) and attribute that pattern to feedback from neighbors to targets. Here, we show that these findings can be reconciled by distinguishing the effects of PND vs. auditory confusability, and by considering the different temporal demands of single-word naming vs. conversational speech.

Phi density is a measure of the auditory confusability of word pairs, based on phone confusion (Iverson et al., 1998; Strand & Sommers, 2011). We calculated the Phi density, i.e. the summed auditory confusability of each word with every other word, for a lexicon of 1,300 CVC words. We fitted linear regression models predicting ND from confusability and vice versa, to identify the variability in each measure not attributable to the other. Decorrelated measures of PND and confusability, were then entered as fixed effects in mixed-effects regression models of vowel dispersion in data from single-word naming and conversational speech, along with control variables, such as frequency and speaking rate.

The outcomes of the mixed-effects models indicate a robust role of auditory confusability, but not neighborhood size (PND) in single-word naming, and a robust role of neighborhood size, but not auditory confusability, in conversational speech: In **study 1**, we model vowel dispersion in the recordings analyzed in Wright (2004), i.e. recordings of two sets of words, termed 'easy' (high frequency, low PND) and 'hard' (low frequency, high PND). We show that Phi density is associated with greater vowel dispersion (p <.01) and produces significant model improvement (χ^2 =5.9, p = .02), whereas adding PND to the model fails to improve model fit significantly. In **study 2,** we model the data analyzed in Gahl et al. (2012). We show that, whereas PND is associated with reduced vowel dispersion and significantly improved model fit (χ^2 = 4.3, p = .03), confusability fails to improve the model significantly (χ^2 = 1.32, p = .25).

Taken together, the results indicate that some effects attributed to PND (the conversational speech data) are likely due to PND, as previously thought, whereas others (the single-word naming data) are due to auditory confusability, rather than PND. We discuss our findings in light of (a) temporal demands of word naming vs. conversational speech, and (b) models of competition in the lexicon (Chen & Mirman, 2012).

Chen, Q., & Mirman, D. (2012). Competition and cooperation among similar representations. *Psychological review, 119*(2), 417-430.
Gahl, S., Yao, Y., & Johnson, K. (2012). Why reduce? *JML, 66*(4), 789-806.
Iverson, P., Bernstein, L., & Auer, E. (1998). Modeling the interaction of phonemic intelligibility and lexical structure in audiovisual word recognition. *Speech Comm, 26*, 45-63.
Luce, P. A., & Pisoni, D. B. (1998). Recognizing Spoken Words: The Neighborhood Activation Model. *Ear & Hearing, 19*(1), 1-36.
Pitt, M. A., et al. (2007). Buckeye Corpus of Conversational Speech (2nd release).
Strand, J., & Sommers, M. (2011). Sizing up the competition: Quantifying the influence of the mental lexicon on spoken word recognition. *JASA, 130*(3), 1663-1672.
Wright, R. (2004). Factors of lexical competition in vowel articulation. *Papers in Laboratory Phonology VI* (pp. 26-50). Cambridge: Cambridge University Press.

Poster Session 2

How modular is lexical category disambiguation?

Peter Baumann (Northwestern University)
baumann@u.northwestern.edu

Language processing requires the integration of bottom-up information and top-down expectations. How (and when) this integration happens distinguishes different architectures: constraint-satisfaction models consider all available information simultaneously, while more modular/sequential architectures use relevant information only when needed. One particular process, for which these two types of architecture make contradicting assumptions, is lexical category disambiguation: constraint-satisfaction models assume that rich contextual information is utilized to determine the category of a word, while modular/sequential models rely on context-independent statistics. Previous research has provided evidence for both architectures: on the one hand, Boland & Blodgett (2001) showed that readers are sensitive to the relative frequencies of lexical categories for a word independent of the context it appears in. On the other hand, Juliano & Tanenhaus (1993) report evidence for some context-dependence in disambiguating the determiner vs. complementizer reading of *that*. Corley & Crocker (2000, C&C) showed that both results can be explained by a model, which incorporates two bottom-up statistics: the context-independent probability of a word given a lexical category (*word-category probability*) and the probability of a lexical category given the lexical category of the previous word (*category co-occurrence probability*). Since both of these probabilities can be calculated independently of context or any further syntactic processing, C&C argue for the existence of a module for lexical category disambiguation.

The present work evaluates C&C's model as a predictor for reading times in the Dundee corpus (Kennedy & Pynte, 2005), showing that readers are sensitive to word-category probabilities in reading natural texts, and that category co-occurrence probabilities (category bigrams) explain variance in reading data in addition to syntactic top-down expectations and word co-occurrence probabilities (word bigrams). We approximated lexical categories by part-of-speech (PoS) tags obtained from the CLAWS tagger (Garside, 1987), and estimated word-category probabilities and category co-occurrence probabilities from a combination of the Dundee Corpus and the British National Corpus. First-pass reading times were modeled using linear mixed-effects regressions with subjects, word tokens and text number as random effects. We first fitted a baseline model with known predictors of reading times: word position, word length, frequency of current and previous word, and word co-occurrence (bigram) probability. Adding word-category probability or category co-occurrence probability as a predictor yields a reliably better model fit (lower AIC and BIC, higher log-likelihood) than the baseline, and a model with both predictors significantly improves over the baseline and over models with only one of them. To test whether the effects of our two predictors may be accounted for by context-dependent syntactic expectations we calculated surprisal (Hale, 2001) based on a non-lexicalized PCFG and added it to the baseline model. Adding word-category probability and/ or category co-occurrence probability as a predictor still yields significantly better model fits than the baseline with surprisal.

Our findings show that both bottom-up statistics in C&C's model are significant predictors of reading times in natural texts, and thus provide further evidence for modular/sequential models relying on simple context-independent statistics for lexical category disambiguation. The observation that category co-occurrence probability is a significant predictor of reading times, even if controlled for word co-occurrence probability and surprisal, shows that syntactic top-down expectations and lexical bottom-up information make independent contributions to reading times, which may be attributed to different comprehension processes or modules.

Speaker distraction interrupts prosodic cues to discourse status

Jennifer E. Arnold, Giulia C. Pancani & Elise Rosa (UNC Chapel Hill)
jarnold@email.unc.edu

Prosody can be ambiguous. It can indicate information status, such that accented words tend to be used when the referent is discourse new, while reduced words are reserved for given information. But prosody can also reflect production difficulty, e.g. hesitant speech may reflect distraction. Although hesitation and accenting do not sound identical, difficulty results in longer words, and lengthening is one cue to accenting. Thus, length can contribute to the perception of both accenting and disfluency. Furthermore, listeners are sensitive to both accenting as an indicator of information status (Dahan et al., 2002, *JML*), and disfluency as an indicator of production difficulty (Arnold et al., 2007; *JEP:LMC*). This raises questions about how speakers interpret lengthening in different discourse contexts.

Here we ask how listeners respond to utterances produced by a fluent speaker vs. a distracted speaker who hesitates and speaks slowly overall. One possibility is that the information-status function of speech is primary, and listeners calculate the relative prominence of words in a distracted context. That is, listeners might know that a slow-down due to distraction alone does not reflect discourse-new information status unless the target is even more prominent than the context. Another possibility is that distracted speech increases the acoustic prominence of all words, even those that are relatively reduced, leading to a temporary preference for new referents during on-line word recognition.

In two visual world eyetracking experiments, listeners responded to instructions like Put the bagel on the circle. Now put the {bacon/BACON/ bagel/BAGEL} on the square. The target was the object in sentence 2, which was either given or new, and either acoustically prominent or reduced. All subjects heard two blocks: a) fluent, and b), one in which the speaker was supposedly distracted by a secondary task. In this block, instructions were spoken with longer word durations and hesitant pauses. Prominence was achieved with higher pitch in both blocks. In Exp. 1, the instructions were recorded naturalistically, and duration correlated with acoustic prominence only in the fluent block. In Exp. 2, the fluent targets were manipulated in Praat to produce longer distracted-sounding targets that were otherwise acoustically matched.

Exp. 1: In the fluent condition, we found the expected effect of acoustic prominence: reduced targets led to faster given target looks, while prominent targets led to faster new target looks. By contrast, the distracted condition yielded little effect of target word prominence, and instead there was a consistent new bias. This result is similar to the reported discourse-new bias for disfluent speech (Arnold et al., 2004, *Psych. Science*), except that in the current experiment, speech difficulty reflected general distraction, not discourse newness itself. Additional analyses revealed that in the fluent context, the preference for the given targets was supported by short duration and low pitch. By contrast, in the distracted context only pitch variation modulated the new bias, and not duration. Exp. 2 yielded the same general new bias in the distracted condition, although the increased pitch cues led to a small relative effect of prominence. These findings demonstrate the ambiguity inherent in prosodic cues, which can signal either discourse status or speech difficulty. Listeners are sensitive to both; yet they occur as part of the same signal, and speech difficulty can interfere with prosodic cues to information status, possibly by producing cues that masquerade as discourse-new acoustic prominence
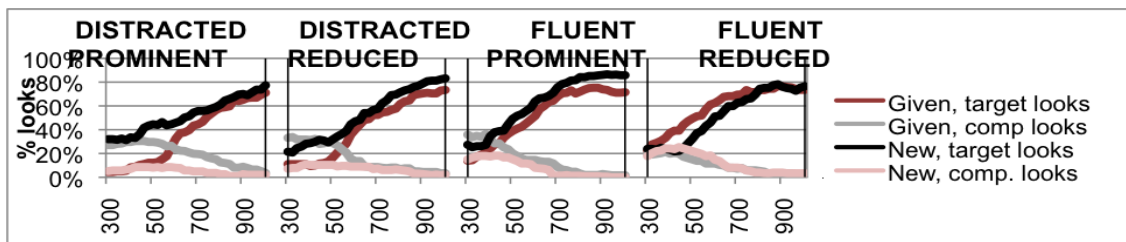

Figure 1. Exp. 1: Looks to targets and competitors from 300-1000 msec after target onset

**Consequences of 'music to one's ears': Structural integration priming from music to language**

Mythili Menon, Elsi Kaiser (University of Southern California)
mythilim@usc.edu

Introduction. Although there have been theoretical claims regarding the relationship between language and music (Jackendoff'06, Patel'08, Katz and Pesetsky'11) few psycholinguistics studies have looked into the question of whether music can prime language – in particular, whether structural aspects in music can prime structural aspects of language. If music and language have overlapping resources, there should be priming from music to language. Initial evidence comes from Van de Cavey/Hartsuiker'11, who used sequences of tones of sine frequencies with different pitch levels, and found priming effects on relative clause (RC) attachment in Dutch.

Study aims. We investigated **(i)** whether musical notes played on *piano* (stimuli that are easily recognized by listeners as being part of the *domain of music*) also trigger priming in the linguistic domain, and **(ii)** whether potential priming effects are strengthened by pauses (given that RC attachment is sensitive to prosodic information). Testing English, which has a low attachment (LA) RC bias–unlike Dutch with a high attachment (HA) bias (e.g. Mitchell/Brysbaert'98)–also allows us to check for potential asymmetries in the 'primability' of LA vs. HA languages. Our participants completed RC fragments after listening to musical melodies.

Methods. We used a 2x2 design with (a) **attachment height** (high attachment/low attachment) and (b) **pause** (pause/no-pause) as factors. We primed participants (N= 20, 30 targets, 120 fillers) with musical melodies created using the 'Circle of Fifths' where tones from musical keys combine to form permissible chords and chord progressions. **Musical primes:** Participants listened to melodies consisting of eight notes (on piano). In high-attachment (HA) primes, the last two notes constituted a transition back to the initial harmonic domain. In low-attachment (LA) primes, there was only an initial part and a second part; the last two notes attached back to the second harmonic domain (see figure). The durations and intervals of the notes were constant, except that in the Pause conditions, there was a pause before the last two notes (No-Pause conditions had no pause). **Linguistic targets:** People completed RC fragments e.g. "*Kevin counted the fans of the singer who*" (position of plural/singular nouns was counterbalanced). We analyzed whether continuations modify the complex NP "*the fans of the singer"* (high attachment) or the lower NP "*the singer"* (low attachment).
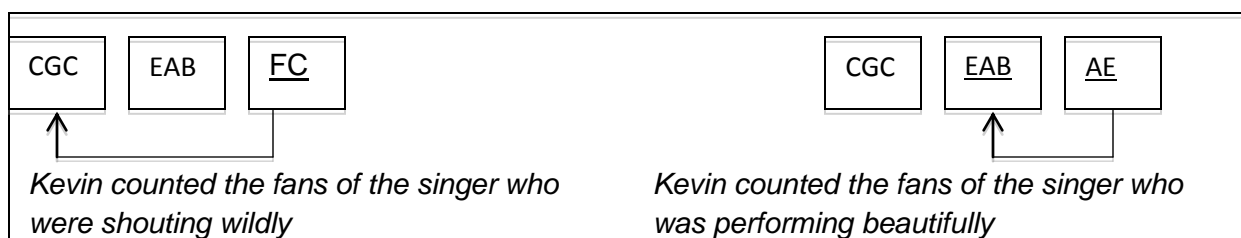
Results. We find significant effects of priming from musical stimuli: Musical HA primes resulted in significantly more HA continuations (38 % HA on average) than musical LA primes (23% HA on average, p<0.01). There was no effect of pause/no-pause and no interaction (p's>.3). Our results show that explicitly musical stimuli (piano notes) activate hierarchical representations that overlap with the syntactic representations of language on an abstract level. Furthermore, pauses after the second noun do not strengthen the HA bias (cf. Fodor'98 on RC prosodic phrasing). This shows that rhythmic cues are not needed to create the priming effect, indicating that what matters is the structure of the harmonic domains.

Conclusion. Our findings suggest that language and music share abstract representations, compatible with the claims of Patel'03, Patel'08. The results of our experiment provide striking evidence for the domain-general level of abstraction in the representation of hierarchical structural information.

Rapid adaptation in the pragmatic interpretation of contrastive prosody

Chigusa Kurumada (Stanford University), Meredith Brown & Michael K. Tanenhaus (University of Rochester)
kurumada@stanford.edu

The realization of prosody varies across speakers, accents, and speech conditions. Listeners must navigate this variability to systematically converge on prosodic interpretations. However, the cognitive mechanisms for this process have only begun to be investigated (Dennison & Schafer, 2010; Watson, Gunlogson, & Tanenhaus, 2008). Building on studies on phoneme perception (e.g., Maye et al., 2002; Norris et al., 2003), we propose that listeners rapidly adapt to speaker-specific prosodic cues in accordance with distributional properties of the input.

Experiment 1 investigated whether listeners rapidly adapt their prosodic comprehension to the distribution of prosodic patterns in the input. We used the English construction "It looks like an X" pronounced either with 1) nuclear accent on the final noun (**Noun-focus prosody**) or 2) L+H* on the verb "looks" (**Verb-focus prosody**). Noun-focus prosody indicates X as the identity of a referent while Verb-focus prosody implies otherwise by evoking contrastive inference (e.g., *It LOOKS like a zebra but it is not one*). We resynthesized 12-step continua of "It looks like an X" sentences, gradually shifting f0 and constituent duration from typical Noun-focus values to typical Verb-focus values. Continua were then normed to determine the maximally ambiguous stimulus (i.e., for which hearers were least certain about the intended meaning).

Subjects were randomly assigned to one of two groups (Noun-focus bias vs. Verb-focus bias conditions). In the training phase, both groups completed a 2AFC judgment task (30 trials), receiving disambiguating feedback after each response. Subjects in the Noun-focus bias condition heard 15 prototypical Verb-focus sentences and 15 instances of the acoustically ambiguous item, disambiguated as Noun-focus ("… and it is one"). Subjects in the Verb-focus bias condition heard 15 prototypical Noun-focus sentences and 15 acoustically ambiguous items, disambiguated as Verb-focus ("… but it is not"). In the following test phase (12 new items, no feedback), subjects in the Verb-focus bias condition more frequently interpreted new sentences as "it is not an X": Subjects shifted their prosodic interpretations according to the distributional information in the input.

Experiment 2 tested whether prosodic adaptation is sensitive to the reliability of prosodic cues. We constructed two experimental conditions in which subjects received different distributions of feedback. In one, subjects heard Noun-focus and Verb-focus prosody mapped onto two different interpretations categorically (High reliability); in the other, the mapping was only probabilistic (Low reliability). We found that lower reliability of prosodic cues resulted in down-weighting of prosody: Subjects were significantly less likely to interpret Verb-focus prosody as "It is not an X". This closely resembles findings from phonemic adaptation, where listeners continuously integrate relevant acoustic cues (e.g., VOT for distinguishing /b/ and /p/) and adjust their weight according to their reliability. Unreliable cues or ad-hoc mispronunciations are down-weighted in the learning process (e.g., Clayards et al., 2008; Kraljic et al., 2008).

To best reflect the speaker-specific distributional patterns, listeners rapidly shift their interpretation of prosodic cues according to (i) how likely they are to signal contrastive meaning; and (ii) how reliable they are. This constitutes novel evidence, beyond phoneme perception, that listeners maintain flexible mapping between variable acoustic signals and linguistic representations through rapid adaptation.

Predicting upcoming words but not semantic features: Evidence from ERPs

Nayoung Kwon (Konkuk University), Pan Liu (Nanyang Technological University) & Patrick Sturt (University of Edinburgh)
nayoung.kw@gmail.com

Previous studies have shown that people build expectation of specific upcoming words based on preceding context (DeLong et al., 2005; Van Berkum et al., 2005). Thus, for example, N400 was elicited to an unexpected article "an" when the predicted word was "kite". However, it is not yet clear whether this pre-activation is only form-based or some levels of semantic features are also pre-activated (cf. Federmeier & Kutas, 1999). To answer this question, we conducted an ERP experiment with sentences with classifiers in Chinese.

Three experimental conditions with congruous, related and incongruous classifiers and head nouns (1) were constructed based on a sentence-completion study (n=30 native Mandarin speakers). Cloze probability of classifiers and head nouns was .68 and .85 respectively in the Congruous condition, and was all 0 in the Related and Incongruous conditions. Head nouns of the Related condition were decided based on two criteria: 1. overlap in semantic features with head nouns of the Congruous condition (e.g., thematically/functionally related), 2. incompatibility of their classifiers with head nouns of the Congruous condition. The intended semantic distance between classifiers in the three conditions was confirmed in a follow-up norming study (n=10 native Mandarin speakers).

If readers predict a specific head noun based on preceding context, N400 will be elicited to classifiers in the Incongruous and Related conditions, as these classifiers are not compatible with predicted nouns. However, if semantic features of head nouns are all pre-activated, the amplitude of N400 to the Related condition will be reduced compared to that of N400 to the Incongruous condition, as the classifiers of Congruous and Related conditions are semantically more related (Federmeier & Kutas, 1999).

14 native speakers of Mandarin participated in an ERP experiment (word-by-word presentation, SOA 800 ms, ISI 400 ms; 64 channels). In a latency window of 300-500 ms post-stimulus onset of classifiers, both the Incongruous [$F(1, 13)=9.11$, $p < .01$] and Related [$F(1, 13)=4.22$, $p < .061$] conditions elicited a N400-like negativity compared to the Congruous condition. The Incongruous and Related conditions did not differ from each other in the N400 time window. However, in a later time window (600-800ms post classifier onset), the Incongruous condition elicited greater negativity compared to the Related condition, which now patterned with the Congruous condition. The N400 pattern at the head noun was similar to that at the classifier, with no evidence for a difference between related and incongruous conditions.

The N400 responses to classifiers of the Incongruous and Related conditions are consistent with previous findings that language users predict a specific upcoming word (DeLong et al., 2005; Van Berkum et al., 2005). Thus, when classifiers are not compatible with the predicted head noun, an N400 is elicited. On the other hand, the lack of difference in the amplitude of N400 to classifiers of the Related and the Incongruous conditions suggests that pre-activation of semantic features is limited, although late positivity effect at the classifier may indicate that semantic features are used in the integration of the classifier.

(1)　邮箱里的　　　　　　　　那　　　(封/份/滴)　　　(信/稿/血)　　　　　　　是从国外寄来的。
　　　which is in a mailbox　　that　　(classifier)　　(letter/draft/blood)　　came from overseas
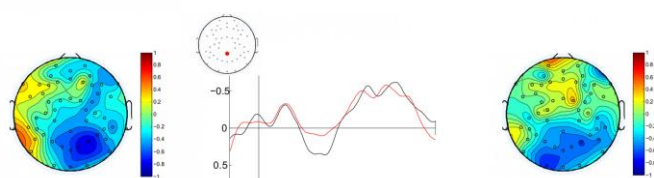　　　'That letter[cong]/draft[related]/blood[incong] which is in the mailbox came from overseas'

Black line: Incongruous−Congruous
Red line: Related−Congruous
Left topographic scalp map:
Incongruous−Congruous; 300-500ms
Right topographic scalp map:
Related−Congruous; 300-500ms

Figure 1 Difference waves and Topographic maps

Self or other: Interplay of verb biases and syntactic constraints during reflexive processing

Xiao He & Elsi Kaiser (University of Southern California)
hex@usc.edu

We report an eye-tracking study investigating effects of verb semantics on the binding-theoretic constraints on the processing of the Chinese reflexive *ziji* "self". According to [1], Chinese referents' accessibility depends on intervening referents' person features: Intervening third-person referents allow *ziji* to refer to the local or long-distance (LD) subject (ex.1a&c); but first-person interveners block LD binding (ex.1b). However, [3] suggest that blocking can be overridden by self-vs.-other-oriented verb biases.

We used visual-world eye-tracking to (i) test when/how intervening subjects' person-features modulate LD referents' accessibility in real-time. Furthermore, given that researchers disagree on whether comprehenders integrate verb semantics predictively or only upon encountering anaphor [2], we also (ii) examined how/when verb biases are integrated during processing to modulate the blocking effect.

1a.	**wo$_1$**	shuo	**Bill$_2$**	zai gongyuan	**fanxing/jianshi**-le	**ziji$_{1,2}$**
	I$_2$	say	**Bill$_2$**	at park	**reflect-upon/spy-on**-perf	**SELF$_{1,2}$**
1b.	~~**John$_1$**~~	shuo	**wo$_2$**	zai gongyuan	**fanxing/jianshi**-le	**ziji$_{*1,2}$** ←BLOCKING
	~~**John$_1$**~~	say	**I$_2$**	at park	**reflect-upon/spy-on**-perf	**SELF$_{*1,2}$**
1c.	**John$_1$**	shuo	**Bill$_2$**	zai gongyuan	**fanxing/jianshi**-le	**ziji$_{1,2}$**
	John$_1$	say	**Bill$_2$**	at park	**reflect-upon/spy-on**-perf	**SELF$_{1,2}$**

Methods. We manipulated (i) the other- vs.-self-oriented verb biases (cf. English *kick*→other-oriented; *shave* → self-oriented) and (ii) the person features of matrix and embedded subjects (e.g., ex.1a-c). (Biases of Chinese verbs were normed for by [3]). Participants (27) listened to sentences while looking at corresponding pictures, and clicked on the last-mentioned thing. On the targets, the last mention thing is the referent of *ziji*.

Predictions. If Blocking strictly determines antecedent choices, 1st person blockers should render LD subjects *unavailable* regardless of verb bias. But if Blocking is modulated by verb bias, participants should look at and click on LD subjects more often with other-directed than self-directed verbs. Also if verb biases cause listeners to activate more probable referents early on, participants should look at subjects consistent with verb biases even before they fully process the reflexive (within first 200 *ms* of the onset of *ziji*, or earlier).

Results. Blocking trials with other-directed verbs showed a higher-than-expected rate of LD choices than blocking trials with self-directed verbs (64% *vs.*10%, p <.0001). But the rate of LD choices on blocking trials with other-directed verbs was lower than on non-blocking trials with other-directed verbs (64% *vs.* 81% & 90%; p's < .01). This suggests that–contra [3]–verb bias undermines but does not fully override blocking. Eye-movements show an early sensitivity to **self-directed** verb bias: After self-directed verbs, people look more at local subjects than LD subjects (p's<.01). This bias starts at the onset of *ziji* (i.e., eye-movements programmed prior to *ziji*). Sensitivity to **other-directed** verbs (i.e., preference to look at LD subjects) emerges at 200*ms* after *ziji*, even on blocking trials (p's<.05). But this effect was modulated by the person-features of LD and local subjects: Listeners looked at LD subjects more in the 3rd-3rd condition than the 1st-3rd/3rd-1st condition (p's<.05). Thus, a person-feature mismatch between LD and local subjects reduced LD subject accessibility. Thus, the on-line effects of blocking may be symmetric, impacting both the 3rd-1st condition and the 1st-3rd condition.

In sum, blocking is not absolute and can be modulated - though not overridden - by **verb semantics**. Moreover eye-movements show that self-vs.-other verb biases are incrementally integrated in real-time: Listeners immediately use verb biases to activate referents in the discourse even before they fully process the reflexive.

References. [1] Huang/Liu'01. Logophoricity, attitudes and ziji at the interface. **[2]** Järvikivi/Pyykkönen'10. Activation and persistence of implicit causality information in spoken language comprehension **[3]** Schumacher et al'11. Perspective in the processing of the Chinese reflexive *ziji*: ERP evidence.

Expectation adaptation for clustering of syntactic structures

Mark Myslín & Roger Levy (UC San Diego)
mmyslin@ucsd.edu

Syntactic processing is affected by expectations of upcoming structures (Hale, 2001; Levy, 2008), and these expectations adapt quickly to recent experience (Fine & Jaeger, 2011; Fine et al., 2010). However, it is not clear whether they are sensitive to higher-order contingencies beyond the individual sentence level, such as the potential for a structure's occurrences to cluster within discourses. We show evidence that comprehenders indeed track longer-range distributions capturing sequences of syntactic structures. Expectations at this level are consistent with adaptation effects both at and beyond the individual sentence level.

We investigate comprehender sensitivity to higher-order contingencies between structures using the sentential complement (SC) construction, which participates in a temporary ambiguity with the direct object (DO) construction (2a, 2c) (Garnsey et al., 1997). Comprehenders quickly adapt their expectations based on recent experience: SC processing speed increases with number of SCs (vs. DOs) experienced (Fine & Jaeger, 2011; Fine et al., 2010). This is a reasonable behavior if comprehenders are sensitive to contingencies beyond the sentence level, since SC tokens cluster in natural language: in the Brown corpus, given an SC, the probability that the following sentence contains another SC (0.11) far exceeds chance (0.04) (significant according to a binomial test: $p < 0.001$). Since comprehenders track these contingencies between structures across sentences, it could also be adaptive to modulate expectations accordingly, as well as to learn the particular distributions of between-sentence contingencies in the current discourse on the basis of recent experience.

We tested this prediction in a between-participants self-paced reading experiment in which two groups ($n = 64*2$) were trained on two-sentence vignettes with different SC clustering properties. In the *clustering* training condition, all SCs occur in clusters of two, as in (2a). In the *anti-clustering* training condition, half of all SCs follow a non-SC structure, and half are followed by a non-SC structure, as in (2b). Both conditions contain identical numbers of SC sentences and the same total number of items (with SC-free vignettes making up the difference in the clustering condition). In half of the SCs, a potential temporary DO ambiguity was blocked by complementizer *that,* and in half it was not. Following training, both groups were tested reading SC.SC vignettes (2a) and SC.DO vignettes (2c). The critical region begins at the disambiguating word in the second sentence (*was* in 2a, *very* in 2c) and ends at the end of the vignette.

If comprehenders modulate expectations based on higher-order, between-sentence contingencies based on recent experience, participants trained in the clustering condition should especially strongly expect SCs as the second sentences of test vignettes, which all begin with SC sentences. We thus predicted an interaction of training condition and continuation of second sentence, such that cluster-trained participants read second sentences especially rapidly when they were SCs. This interaction was significant in a mixed-effects model of residual reading times in the critical region ($t = -2.10$, $pMCMC < 0.05$), indicating that comprehenders do learn between-sentence contingencies based on recent experience.

Syntactic adaptation effects both at the individual sentence level and across multiple sentences are consistent with an account in which comprehenders track longer-range distributions capturing sequences of syntactic structures. Speeded processing for individual structures given recent experience with the structure is consistent with an expectation for clustering of tokens, a reasonable prior expectation given existing distributions in natural language. We show such expectations are rapidly adaptive, shaped by the particular higher-order, discourse-level distribution of contingencies between structures that is experienced.

(1) Ruby confided (that) her secret had been really bothering her for a long time.
   (2a) Her friend whispered (that) the only solution was to dispose of the evidence (SC.SC)
   (2b) Her friend was always willing to listen to her problems and be there for her. (SC.Other)
   (2c) Her friend whispered the only solution very quietly in her ear and hugged her. (SC.DO)

How speakers trade accuracy for speed when producing subject-verb agreement

Laurel Brehm & J. Kathryn Bock (University of Illinois)
lbrehm3@illinois.edu

Trouble in language production sometimes surfaces in errors and delays. Since these two symptoms of difficulty can trade off, theories may make predictions that are confirmed with measures of accuracy but disconfirmed with measures of speed, and vice-versa. In work on grammatical agreement, variations in the accuracy of production have been found that are consistent with certain accounts of agreement mechanisms, while variations in speed are at odds with the same accounts. To reconcile these contradictions, we carried out three experiments that gauged speed and accuracy in producing agreement. The data were analyzed in a statistical model that integrates the measures into a coherent framework involving agreement mechanisms that are sensitive to separable conceptual and linguistic forces.

All of the experiments assessed the impact of factors associated with the notional and grammatical number determinants of agreement. The notional factors included conceptual integration and abstractness, which create differences in referential cohesion. The grammatical factor was the singularity or plurality of distractor (non-head) nouns. These grammatical number variations are important in the creation of attraction, in which a verb appears to erroneously agree in number with a distractor rather than the head, as in *The label on the bottles **were** ugly* [1]. In these cases, agreement variation could stem from either the effect of ambiguous construals of the subject's numerosity (e.g. of *label* as a type or multiple tokens of a type), or from the grammatical number of the distractor (*bottles*). Some accounts treat these as separable factors in agreement; others regard them as interacting components of agreement. Tests of production speed have supported the former view [2] and tests of accuracy the latter [3].

In all experiments, speakers completed a visually or aurally presented sentence preamble (e.g. *The drawing of the flowers*) with a verb and predicate (e.g. "was bad"). Variations in verb number (singular/plural) and production latency (from offset of preamble to onset of verb) were measured and mathematically modeled with a technique that treats accuracy and speed together (EZ-diffusion [4]).

EZ-diffusion provides four parameters that characterize the processes that enter into the production of agreement. One parameter (T_{er}) corresponds to the speed of sentence subject conceptualization and other pre-decision planning. This parameter was highly sensitive to difficulty in construing notional number, and less sensitive to grammatical number. Another parameter (v), corresponding to the difficulty of formulating subject *number*, was sensitive to notional and grammatical sources of plurality, including referential coherence, relational plausibility within the sentence, and the grammatical number of distractors. A third parameter (*MDT*) corresponds to the speed of the verb number decision, and was sensitive to conflicting grammatical number and sources of lexical-semantic interference, but not to notional factors. The fourth parameter (a) reflects the criterion for the agreement decision: How conservative is the speaker in selecting the verb form? This parameter was sensitive to grammatical number, with mismatching local nouns prompting less conservative responses.

The model's outcomes support the conclusion that the mechanisms of agreement are differentially sensitive to two factors: Notional number affected some parameters and grammatical number affected others. The results thus put speed and accuracy together into a coherent picture of the production of subject-verb agreement, offering a sharper perspective on the how and when of language production.

References. [1] Bock, K., & Miller, C.A. (1991). Broken agreement, *Cognitive Psychology*, 23(1), 45-93. [2] Brehm, L., & Bock, K., (2012). What counts in grammatical number agreement? (Submitted). [3] Solomon, E. S., & Pearlmutter, N. J. (2004). Semantic integration and syntactic planning in language production. *Cognitive Psychology,* 49, 1-46. [4] Wagenmakers, E.J., van der Maas, H.L. & Grasman, R.P.P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review,* 14, 3-22.

Poster Session 2

Recent experience changes production preferences in the face of semantic biases

Victor Ferreira & Liane Wardlow (University of California, San Diego)
vferreira@ucsd.edu

Speakers can often express ideas with different structural alternatives, such as prepositional datives (*the doctor sent the notice to the trainer*) versus double objects (*the doctor sent the trainer the notice*). However, biases can lead to preferences or dispreferences for particular options. One type of bias arises because structure-production mechanisms themselves have preferences, which can change based on recent experience. For example, after exposure to many double objects, subjects are more likely to produce more double objects in the future (Kaschak, 2007). The malleability of speakers' preferences as a function of recent experience is likely advantageous, for 'alignment' of production among conversational partners (Pickering & Garrod, 2004), and for adaptation when our linguistic environment changes (e.g., because of geographical or demographic shifts). Another type of bias is semantically based: With inanimate recipients, double objects are dispreferred (*the doctor sent the notice to the store* vs ?*the doctor sent the store the notice*). Here, we pit the adaptability of structural biases as a function of recent experience against the semantically based bias against inanimate-object double-objects. Are semantically based biases immutable (perhaps because they are inherent to the syntactic construction or to production processes)? Or can they be overcome by recent experience? If so, must the recent experience violate the semantically based bias itself?

The experiments proceeded through pre-test, exposure, and post-test phases. Pre- and post-tests were identical: Subjects were given two postverbal noun phrases and then a subject and verb, and asked to form full sentences. Critical test phrases elicited inanimate recipient sentences (*the notice*, *the store*). In Experiment 1, the exposure phase included 36 inanimate recipient sentences. For (about) half of subjects, these were (acceptable) prepositional datives, whereas for the other half, these were (dispreferred) double objects. If the dispreference for inanimate-recipient double-objects can be overcome at all, then experience with inanimate-recipient double-objects should increase inanimate-recipient double objects on post-test relative to pre-test. Indeed, after exposure to (dispreferred) inanimate-recipient double objects, a 10% increase in inanimate-recipient double-object productions was observed, but after exposure to (acceptable) inanimate-recipient prepositional datives, production preferences were unchanged. Thus, the dispreference for inanimate-recipient double-objects can be overcome.

But perhaps experience with *inanimate-recipient* double-object sentences specifically is not required to overcome this dispreference. Experiment 2 assessed this by changing exposure such that all 36 exposure sentences were double-objects. For (about) half of subjects, these had inanimate recipients (as in Experiment 1), but for the other half, these had animate recipients. If the dispreference can be overcome by double-object sentences generally, then even subjects exposed acceptable *animate*-recipient double objects should show increases in inanimate-recipient double-object production. Indeed, this was observed: Subjects exposed both to dispreferred inanimate-recipient and acceptable animate-recipient double objects showed significant and equal (5-8%) increases in inanimate-recipient double-object production post-test.

In all, these results show that recent linguistic experience can overcome meaning-based biases against the production of particular structural options, and that the experience need not be specific to the meaning-based bias itself. The malleability imparted to the language production system through adaptation from recent experience trumps semantically based biases against the production of particular structural options.

Advance planning of verbs in head-final language production

Shota Momma, Robert Slevc, & Colin Phillips (University of Maryland)
shotam@umd.edu

Given their central role in determining structures of sentences, one might expect verb selection to occur early in the production of verb-final languages, but previous evidence suggests otherwise (Schriefers et al. 1998). Here we resolve this puzzle based on results from 3 picture-word interference (PWI) studies in Japanese. The results show that verbs are selected ahead of when they must be uttered, but not as early as the beginning of a sentence.

The temporal locus of verb selection is controversial (cf. Griffin and Ferreira 2006). One plausible assumption is that verbs are selected before phonological encoding of the subject (or first) NP is finalized (Ferreira 2000). However, a series of PWI studies in German (Schriefers et al. 1998) shows that verb-related distractors interfere with speech onset only in constructions requiring verb-initial order, suggesting late selection of verbs. However, Schriefers' study may have overlooked evidence of early verb selection by probing too early in the German sentences. Linguistic observations suggest that verbs are more strongly tied to their objects than their subjects, both syntactically (w.r.t. case assignment and constituency) and semantically (w.r.t. argument status; Kratzer 1996). We therefore tested whether verb selection might occur before phonological coding for object nouns is finalized, but not before utterance of a sentence-initial subject noun. Japanese is well-suited to test this hypothesis, as it is a strictly head-final language that allows liberal argument dropping. These features allowed us to elicit complete, well-formed SV (intransitive) and OV (transitive) sentences, enabling us to measure speech onset of subject-initial and object-initial sentences without changing the position of verbs. We tested the timing of verb selection via 3 PWI studies.

The same set of native Japanese speakers (n = 20) participated in all three experiments. They were asked to produce a verb (Exp. 1), a sentence (Exp. 2) or just a noun (Exp. 3) corresponding to the target pictures, while ignoring superimposed distractor words. The results are summarized in the figure below. Exp. 1 was designed to confirm that the distractor verbs were suitable for eliciting semantic interference (SI) effects, in the form of delayed speech onset times, when the distractor preceded a verb that was produced in isolation. Comparable SI effects were found with transitive and intransitive verbs, showing that the distractors were suitable. Exp. 2 was designed to test if the SI effect observed in Exp. 1 persists in simple sentence production when subject/object nouns precede the verbs. Using the same set of picture-distractor pairs as in Exp. 1, an SI effect was observed only in the OV sentences. This shows that verbs are selected before the onset of object nouns but not before the onset of subject nouns. Finally, Exp. 3 used a noun-only production task to verify that the selective SI effect in Exp. 2 could not be attributed to an unintended relation between the nouns and the distractor verbs. In this experiment the same noun-distractor pairs used in Exp. 2 yielded no SI effects.
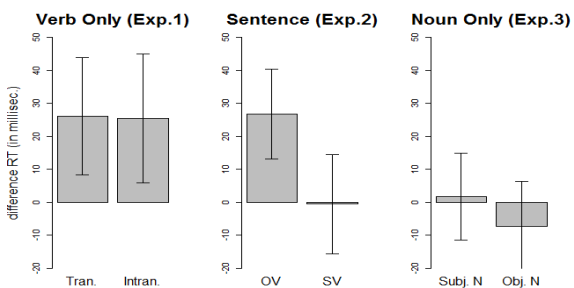
SI effects (related-unrelated) in three experiments. Error bars represent 95% Confidence Intervals.

Taken together, the results indicate that early verb selection does occur in verb-final languages, but that it might occur at the onset of a verb phrase/predicate, rather than at the onset of a sentence. Although some previous studies have shown evidence of advance lexical selection of post-verbal nouns in production of simple phrases (e.g., Meyer, 1996), this study provides the first direct evidence of early verb selection.

Poster Session 2

Silent structures in ellipsis: Evidence from syntactic priming

Ming Xiang, Julian Grove, Jason Merchant, Genna Vegh, Stefan Bartell, Katina Vradelis (University of Chicago)
mxiang@uchicago.edu

When resolving ellipsis, speakers could either derive semantic interpretations by first constructing unpronounced syntactic representations at the ellipsis site[1,2], or they could assign an interpretation to the ellipsis site based on the semantic/discourse information present in the antecedent, without the need to construct any syntax at all [3,4]. Using syntactic priming [5,6], the present study found that syntactic structures are being accessed at the ellipsis site.

During the priming phase, participants first saw a priming sentence consisting of two conjoined clauses, then heard a recording of it, and finally repeated the sentence. The priming sentences follow a 2x3 design (see (1), item N=18; filler N=36). The first clause contained either double-object (DO) or prepositional-dative (PD) structures (the factor of Prime Type); and the second clause was either (i) elliptical, (ii) non-elliptical full structure, or (iii) a neutral intransitive control (the factor of Continuation Type). Following the priming phase, participants then viewed a target picture depicting an unrelated event and were asked to describe it in a full sentence. The depicted event could be described with either a double-object (DO) or a prepositional-dative (PD) structure (e.g. *The girl gave the boy an apple* or *The girl gave an apple to the boy*).

(1) An example of a priming sentence in 6 conditions:
  DO primes: First Ralph sang Sheila a song, and then…
         (a) Marcus did./ (b) Marcus sang her one./ (c) Marcus groaned.
  PD primes: First Ralph sang a song to Sheila, and then…
         (d) Marcus did./ (e) Marcus sang one to her./ (f) Marcus groaned.

We report the proportions of producing PD structures in (2) (subj N=69). Mixed-effect logistic modeling showed a main effect of Prime Type (p<.05, p-value obtained through model comparison), and the interaction between Prime and Continuation type is approaching significance (p=.08). Comparisons between conditions showed a priming effect between the primed and the unprimed non-elliptical full structure (p<.05; 2b vs. 2e, more production of PD under PD prime than DO prime); the priming effect in the ellipsis pair is approaching significance (p<.07; 2a vs. 2d); and there is no difference between the neutral control pair (p>.5; 2c vs. 2f).

(2) Continuation Type	production of PD	
	Unprimed (following DO primes)	Primed (following PD primes)
(i)Elliptical:	(a) 63%	(d) 75%
(ii)Nonelliptical:	(b) 61%	(e) 71%
(iii)Neutral:	(c) 73%	(f) 67%

Data collection is still on-going. Preliminary results above suggest the following. First, elliptical and non-elliptical full structures showed a similar priming effect, suggesting that syntactic representations are being accessed at the ellipsis site, rather than just semantic/discourse representations alone. Previous studies have shown the absence of syntactic priming effect for elliptical structures in Chinese [7], suggesting that elliptical constructions form a heterogeneous class cross-linguistically[8]. Second, considering that (i) there is no overt phonological material at the ellipsis site; and (ii) previous studies have suggested that the syntactic representation at the ellipsis site is retrieved through a "pointer" memory mechanism that directly accesses the antecedent structures [9], rather than being constructed procedurally at the ellipsis site, current data highlights the independent status of abstract syntactic representations in priming.

References:[1]Merchant, 2009. [2]Chung, Ladusaw & McCloskey,1995. [3]Culicover & Jackendoff 2005. [4]Hardt,1993. [5]Bock,1986. [6]Bock & Loebell. 1990. [7]Cai, Pickering, & Sturt, 2012. [8]Merchant&Simpson 2012. [9]Martin & McElree, 2008.

Planning units in Tagalog sentence production: Evidence from eye tracking
Sebastian Sauppe & Elisabeth Norcliffe & Agnieszka E. Konopka (Max Planck Institute for Psycholinguistics) & Robert D. Van Valin, Jr (Max Planck Institute for Psycholinguistics; Heinrich Heine University Düsseldorf; University at Buffalo, The State University of New York) & Stephen C. Levinson (Max Planck Institute for Psycholinguistics; Radboud University Nijmegen)
sebastian.sauppe@mpi.nl

In transforming thoughts into speech, speakers must encode a pre-verbal message linguistically. We investigated this process in the Philippine language Tagalog, which, unlike most European languages, has a verb-initial basic word order.

The Tagalog verb agrees in semantic role with the "privileged syntactic argument" (simplified: "subject"). Either the semantic agent or the semantic patient may be the "subject" (e.g., [1]); in either case the sentence remains transitive (unlike English where promoting the patient to the subject requires passivization). The subject can occur sentence-finally (VOS: [chase$_{patient.voice}$ woman chicken$_{subject}$] – 'the woman chases the chicken') or sentence-medially (VSO: [chase$_{patient.voice}$ chicken$_{subject}$ woman] – 'the woman chases the chicken'), but in both cases, sentence-initial verb position requires selection of appropriate subject–verb agreement marking. We examined the time course of sentence formulation for descriptions of transitive events [2,3] to test whether these grammatical properties entail early planning of the dependency between verb and subject and whether this dependency planning is temporally dissociated from lexical encoding of the subject argument.

Fifty-three native speakers of Tagalog described 44 target pictures depicting two-participant, transitive events interspersed between unrelated filler pictures while their gaze and speech were recorded (Tobii T120 eye-tracker, 120 Hz sampling rate). The distribution of early eye movements (0–600 ms) to the two characters in the target pictures showed that a character was fixated more often when it was selected to be the subject of a given sentence than when it was selected to be the non-subject (the "object"), regardless of whether the subject character was the semantic agent or the semantic patient in the event and regardless of whether the word order was VOS or VSO. This suggests an early phase of sentence formulation that involves the planning of the dependency between verb and subject in order to select the appropriate agreement markers on the verb and that is independent of word order. The extent to which speakers continued fixating the subject character after 600 ms depended on whether the subject immediately followed the verb or was produced sentence-finally: in sentences with VOS word order speakers fixated the post-verbal object character before they re-fixated the subject character, whereas in VSO sentences speakers continued fixating the subject character until shortly after speech onset and then shifted their gaze to the object character. This suggests that sentence formulation in Tagalog can involve two temporally dissociated phases: rapid planning of the dependency relation between verb and subject, driven by agreement marking, followed by lexical encoding of post-verbal arguments in the order of mention. This is in contrast to English, where rigid subject-initial word order conflates the dependency planning and lexical encoding of the subject so that these two phases are not easily separable.

The results speak against a strictly linear view of sentence planning, which assumes that speakers immediately encode a fixated character lexically [2]. Instead, the presence of a grammatical effect of the subject on early eye movements is evidence of linguistic guidance in the earliest stages of message and sentence formulation and suggests that the time course of these processes can be shaped by the structure of the language.

References
[1] Himmelmann (2005). Tagalog. In Adelaar & Himmelmann (Eds), *The Austronesian languages of Asia and Madagascar*.
[2] Gleitman, January, Nappa, & Trueswell (2007). *JML*, *57*, 544–596.
[3] Griffin & Bock (2000). *Psych. Science*, *11*, 274–279.

Poster Session 2

The upside of not having a syntactic choice: Effects of syntactic flexibility on Korean production

Heeju Hwang & Elsi Kaiser (University of Southern California)
heejuhwang@gmail.com

Introduction. Sentence production requires speakers to choose a structural frame, because a message can be often expressed with more than one structure (*syntactic flexibility*, Ferreira 1996). The present study examined how Korean speakers choose between alternative syntactic structures. Two models of sentence production make different predictions regarding the effects of syntactic flexibility. The *competitive model* predicts that syntactic flexibility should cause production difficulties (e.g., longer production latencies and errors because the alternative structures compete for selection, restricting one another's availability. In contrast, the *incremental model* suggests that syntactic flexibility should facilitate production (e.g., shorter production latencies and fewer errors), because it allows more accessible lexical items to be accommodated sooner. To evaluate these models, Ferreira (1996) asked participants to construct sentences using given words. He found that English speakers produced sentences faster and with fewer errors when given syntactic choices (e.g., *I gave/toys/children* vs. *I donate/toys/children*), supporting the *incremental model* for English.

Experiment. Following the methodology of Ferreira (1996), we conducted a production study (n=21) investigating the effects of syntactic flexibility in Korean numeral quantifier (NQ) constructions. Participants were asked to produce sentences using words presented on a computer screen. We manipulated syntactic flexibility via case-particles; when quantifying object nouns, **genitive-marked** NQs (1a) allow speakers to produce only pre-nominal structures (NQ+Noun), whereas **accusative-marked** NQs (1b) allow only post-nominal structures (Noun+NQ) (*non-flexible conditions*). When the NQs are not case-marked (1c), speakers can choose between pre- and post-nominal structures (*flexible condition*). In a separate rating study (n=16), the pre- and post-nominal constructions used in the experiment were confirmed to be *semantically comparable*; the mean comparability rating was 6 on 1-7 scale (1=completely un-interchangeable, 7=completely interchangeable/semantically equivalent).

Mary met seven students:
(Below are words shown on the screen, the locations of words were counterbalanced)
(1a) *Non-flexible/pre-nominal condition***:** Mary-NOM / met / seven-CL-**GEN** ⁄ student
(1b) *Non-flexible/post-nominal condition*: Mary-NOM / met / seven-CL-**ACC** ⁄ student
(1c) *Flexible condition*: Mary-NOM / met / seven-CL ⁄ student

Predictions. We expect that post-nominal constructions should be produced faster and with fewer errors, since they occur more frequently than pre-nominal constructions (Kim & Lee, 2010). Importantly, however, this frequency effect should be *modulated by the presence/absence of syntactic flexibility*. That is, according to the incremental model, both pre-nominal and post-nominal constructions should be produced faster and with fewer errors in the flexible condition than in the non-flexible conditions. But according to the competitive model, the non-flexible conditions should result in shorter utterance latencies and fewer errors.

Results. While the error rates do not provide clear evidence in favor of either the incremental or the competitive model (error rates did not differ significantly between the flexible condition and non-flexible conditions), production latencies point towards the competitive model. As expected, speakers produced pre-nominal constructions significantly more slowly than post-nominal constructions. Crucially, however, both pre-nominal and post-nominal structures were produced *faster in the non-flexible than in the flexible condition* (pre-nominal: 2980ms vs. 3084ms; post-nominal: 2521ms vs. 2642ms, p<.05). This provides evidence for the competitive model for Korean in contrast to the incremental production model for English. We suggest that the different findings in English and Korean are related to how they signal grammatical functions (by word order in English and by case-marking in Korean).

References. Ferreira, V. S. (1996). *Journal of Memory and Language*, 35, 724-755.

Towards the understanding of the correspondence relationship between language-related ERP components and oscillatory activities

Hiroaki Oishi[1], Nobuyuki Jincho[1], & Reiko Mazuka[1,2] ([1]RIKEN Brain Science Institute, [2]Duke University)
oishih@brain.riken.jp

In our previous study, we conducted time-frequency analysis on electroencephalogram (EEG) data to investigate whether the cognitive process underlying the P600 effect elicited by structural revision is the same as that underlying the P600 effect elicited by thematic violation. We recorded EEGs while participants read garden-path sentences and thematically-anomalous sentences in Japanese. We found that both structural revision and thematic violation elicited P600 effects but showed different patterns of oscillatory activities. This indicates that the cognitive processes underlying the superficially identical ERP components can be qualitatively different. However, since the thematic violation also elicited an N400 effect and showed significant power change in a few frequency bands and since little was known about the correspondence relationship between language-related ERP components and oscillatory activities, it remained unclear to which oscillatory activity each type of ERP component was related.

To make the correspondence relationship clear, this study conducted an experiment to obtain brain responses to semantic and syntactic anomalies in Japanese which is predicted to elicit an N400 effect and a P600 effect, respectively (example sentences are shown below). The ERP analyses revealed that the semantically anomalous sentences in which object nouns were implausible THEME for matrix verbs resulted in an N400 effect and that the syntactically anomalous sentences in which verb subcategorization is violated elicited an N400 effect as well as a P600 effect. The time-frequency analyses revealed that both semantically anomalous and syntactically anomalous sentences showed event-related synchronization (ERS) in theta band (4-7 Hz) in 200-600ms time window. Furthermore, syntactically anomalous sentences also showed event-related desynchronization (ERD) in beta band (13-18 Hz) in 200-400ms time window.

Though the N400 enhancement for syntactic anomaly was unexpected, it might reflect the mismatch between expectation and reality with regard to syntactic property of the eliciting word (for a similar finding, see DeLong, Urbach, & Kutas 2005 in *Nature Neuroscience*). The finding that both semantically anomalous and syntactically anomalous sentences elicited an N400 effect and showed ERS in theta band suggests that the cognitive process underlying the N400 effect might be reflected by theta power increase. If this is the case, the beta power decrease observed in syntactically anomalous sentences should be associated with the cognitive process underlying the P600 effect. Interestingly, the oscillatory activity pattern as well as ERP response pattern for syntactically anomalous sentences was similar to those for thematically-violated sentences in our previous study. This suggests that the thematic violation might have been perceived as syntactic anomaly. We will discuss the implications of these findings for models of sentence processing.

(1)   Umaretate-no imomushi-ga midori-no <u>happa-ni</u>/[#]<u>jitensya-ni</u> kajiritsui-ta.   (semantic conditions)

     newborn-Gen caterpillar-Nom green-Gen leaf-Dat/bicycle-Dat bite-Past

   (The newborn caterpillar bit a green leaf/bicycle.)

(2)   John-ga Mary-ni shinsaku-no kuchibeni-o purezento-shita/*tsukaihatashi-ta   (syntactic conditions)

     John-Nom Mary-Dat brand-new-Gen lipstick-Acc give-Past/use up-Past

   (John gave/*used up a brand-new lipstick to Mary.)

Poster Session 2

Are our eyes really faster than our brains? Aligning eye-tracking and ERP time estimates
Wing Yee Chow, Colin Phillips (University of Maryland) & Suiping Wang (South China Normal University)
wychow@umd.edu

A growing body of eye-tracking research demonstrates that readers' are sensitive to high-level syntactic and semantic anomalies within the first 200-250ms [1]. However, similar manipulations often have no effect on event-related potentials (ERPs) until somewhat later [2], leading to debate about the time-alignment between eye-movements and ERPs [3,4]. However, there have been few side-by-side comparisons [4-6], and in these comparisons it may be insufficient to simply use matched materials, due to differences in parafoveal preview, spillover effects, and presentation rate between the paradigms. Differences between eye-tracking and ERPs may reflect sensitivity to different underlying processes, as has been suggested in the case of N400s [3], or they may reflect more task-specific factors. Here we focus on alignment of eye-movements with the P600, a late but widespread ERP component. In a recent study we observed a strong timing discrepancy between eye-movement measures and P600s [7]. Here we present two new eye-tracking studies that more closely match ERP presentation conditions, and show that this leads to close alignment of time-course estimates.

Using the unambiguous S-O-V *BA*-construction in Mandarin Chinese, we tested detection of thematic role-reversal anomalies in sentences with animate pre-verbal arguments (e.g., cop_{SUBJ} $thief_{OBJ}$ arrest vs. $thief_{SUBJ}$ cop_{OBJ} arrest). We previously observed that role-reversals consistently elicited only a P600 effect at the critical word in ERPs, but they led to longer first fixations at the critical word and longer first pass times even in the pre-critical region in eye-tracking [7], suggesting that eye-movements and ERPs might be sensitive to different underlying processes. In the current study we first addressed the possibility that faster presentation rates and delayed oculomotor effects in eye-movements might lead effects of region *n-1* to appear in early measures at region *n*. In Experiment 1 (n=24) we inserted a temporal phrase between the arguments and the verb to create a fully-matched pre-critical region (*a* vs. *b*). We found that first pass reading times in the argument region were longer in the reversed condition than in the control (774ms vs. 732ms; *p* =.03), possibly reflecting sensitivity to argument-role bindings or to ease of predicting an upcoming verb. Crucially, there was no effect of role-reversals at the inserted temporal phrase (*p*s > .10). At the verb, role-reversals increased probability of regression and regression path times relative to controls (22% vs. 15% and 553ms vs. 507ms respectively; *p* < .05), but did not impact earlier measures. The timing of such effects is consistent with the P600 effect in ERPs. Experiment 2 (n=36) used a gaze-contingent boundary paradigm [8] to examine the potential contribution of parafoveal preview, resulting in a 2 (canonical vs. role-reversed) by 2 (valid vs. invalid preview) design. In the invalid preview conditions the target Chinese characters in the critical region were displayed only when the eyes made a saccade to that region. In the critical verb region, first fixations and first pass times were longer in the invalid preview conditions (main effect of preview: all *p*s < .001), but there was no effect of role-reversals. The earliest effect of role-reversals emerged in the post-critical region, where probability of regression was significantly higher in the role-reversed than in the control condition (main effect of role-reversal *p* = .02). The timing of this effect at region *n+1* is compatible with a P600 effect at region *n*. Taken together, these results show that ERP and eye-tracking provide convergent evidence regarding the time course of thematic relation processing.

   a. Canonical: /Cop BA thief/ ZAI yesterday afternoon/ arrest…
   b. Role-reversed: /Thief BA cop/ ZAI yesterday afternoon/ arrest …

References
 [1] Clifton, Staub & Rayner. (2007). In Van Gompel et al. (Eds.) Eye movement research: A window on mind and brain, 341-372. [2] Kutas & Federmeier. (2007). In: Gaskell (Ed.), Oxford Handbook of Psycholinguistics, 385-406. [3] Sereno & Rayner (2003) TiCS. 7, 489-493. [4] Dambacher & Kliegl (2007) Brain Res. 1155, 147-162. [5] Ledoux, Camblin, Swaab & Gordon (2006). Behav Cogn Neurosci Rev. 5, 107-127. [6] Camblin, Gordon & Swaab (2007) JML. 56, 103-128. [7] Chow, Phillips & Wang (2012). CUNY. [8] Rayner (1975) Cog. Psych. 7, 65-81.

Eyetracking evidence for the subject relative advantage in Mandarin

Lena Jäger, Shravan Vasishth (University of Potsdam), Zhong Chen (Cornell University), & Chien-Jer Charles Lin (Indiana University)
lena.jaeger@uni-potsdam.de

In most languages examined so far, subject relatives (SRs) are both more frequent than object relatives (ORs) and also easier to process. This processing advantage of SRs compared to ORs has been explained in terms of the shorter distance between the head noun and its corresponding gap in SRs compared to ORs (Grodner & Gibson, 2005). However, for languages like English, an alternative explanation for the SR advantage is the higher structural frequency of SRs over ORs and hence we cannot tease apart frequency-based explanations from distance-based explanations for the processing difficulties in ORs. A decisive test case is Chinese in which the distance between the head noun and its corresponding gap is shorter in ORs which are less frequent than SRs (Hsiao & Gibson, 2003): for Chinese, frequency-based accounts predict an SR advantage while distance-based accounts predict an OR advantage. There is conflicting evidence on whether SRs or ORs are easier to process in Chinese (e.g., Hsiao & Gibson, 2003; Gibson & Wu, 2013; Chen et al., 2012). A central problem in experiments on Chinese relatives is that there are several local ambiguities that might lead to garden paths confounding the interpretation of the results. Based on the design of our earlier self-paced reading study (Chen et al., 2012), we carried out an eyetracking study (49 subjects, 32 items) by creating a syntactic configuration (see (1)) that ensured that the relative clauses were unambiguous from the outset (in (1), the first three words predict that a relative clause is coming up). We looked at both subject- and object-modifying SRs and ORs.

The results replicated the main findings of our self-paced reading experiment. Overall, we found a main effect of relative clause type across regions and different dependent measures; there was an SR advantage in the relative-clause region (the Noun-Verb/Verb-Noun sequence before the relativizing marker *de*) in regression path duration (t=3.37), total fixation time (t=2.06), right-bounded reading time (t=2.34), re-reading time (t=2.0), first-pass regression probability (z=5.07) and re-reading probability (z=2.5). At the head noun, we observed an SR advantage in total fixation duration (t=2.04) and re-reading time (t=2.5). The SR advantage remained significant in the materials following on the head noun.

1a) Subject relative clause (subject modifying)

Nage zuowan zou le fuwusheng yi dun de **guke** jian guo laoban bingqie jide ta.
Det CL last.night hit ASP waiter one CL REL customer see ASP boss and remember him.
That customer who hit the waiter last night had seen the boss before and remembered him.

1b) Object relative clause (subject modifying)

Nage zuowan fuwusheng zou le yi dun de **guke** jian guo laoban binqgie jide ta.
Det CL last.night waiter hit ASP one CL REL customer see ASP boss and remember him.
That customer who the waiter hit last night had seen the boss before and remembered him.

To conclude, we see clear evidence for a subject relative advantage in Chinese that begins at the relative clause and extends to the head noun and beyond. Thus, the distance-based explanation for relative clause processing is difficult to uphold; the data speak in favor of a structural-frequency based explanation for relative clause processing difficulty.

References. [1] Chen, Jäger, Li & Vasishth (2012). Structural-frequency affects processing cost: Evidence from Chinese relative clauses. In Annual CUNY Sentence Processing Conference 2012, New York, NY. [2] Gibson & Wu (2013). Processing Chinese relative clauses in context. Language and Cognitive Processes, 28(1-2),125-155. [3] Grodner & Gibson (2005). Consequences of the serial nature of linguistic input. Cognitive Science, 29, 261-290. [4] Hsiao & Gibson (2003). Processing relative clauses in Chinese. *Cognition, 90*, 3-27.

Discourse accessibility and structural bias: Processing D-linked phrases in sluices
Jesse A. Harris (Pomona College)
jesse.harris@pomona.edu

Wh-phrases like *which students* in *Which students did you see?* are appropriate only if *D-linked* to context, i.e., if a salient set of students is discourse-accessible (Pesetsky, 1987), and access the discourse representation immediately (Frazier and Clifton, 2002, 2005). This study investigated how the processor resolves globally ambiguous sluice structures with D-linked phrases, e.g. (1), to determine whether the retrieval process was subject to discourse and structural economy constraints simultaneously. Two constraints are proposed to interact within a cue-based, content-addressable retrieval system, in which potential antecedents are accessed in parallel, but linguistic focus guides the focus of attention. First, an immediate structural *Object Bias (OB)* favors antecedents in the object of the matrix clause (Carlson et al, 2009; Poirier et al, 2010) guided by default focus-marking. Second, a discourse-economy constraint, *Alternatives on Demand (AD)*: *when interpreting a D-linked phrase, favor discourse alternatives given by previous discourse*, avoids potentially unnecessary discourse entities during retrieval. Central predictions of the model were tested in six experiments.

Three Internet questionnaires tested the prediction that comprehenders prefer to resolve ambiguous *which* phrases in sluices (1) by obeying both OB and AD. Thus, antecedents that avoid positing discourse-new alternatives, e.g., disjunctions like *Bill or Sue*, should be favored when in object position. Participants (N=32) took the disjunction as the antecedent when in object (1a; M=83%, SE=3) more often than when in subject (1b; M=55%, SE=4) position, z=3.30, p<0.05. A follow up (N=36) replicated the results and confirmed that the effect was not driven by the anaphoric properties of the pronoun *one*. A fill-in-the-blanks task (N=47) substituting a blank for the disjunction (*Bill ___ Sue*) found that participants were more likely to provide a disjunction response for object (M=77%, SE=2) than for subject (M=64%, SE=2) blanks, z=3.77, p<0.001. In offline measures, at least, the processor appears to prefer interpretations that satisfy both OB and AD.

A self-paced reading (N=48) and an eye movements (N=36) study investigated how the processor manages conflict between structural (OB) and discourse (AD) preferences when resolving D-linked phrases during online comprehension. The experiments manipulated, as before, the position of the disjunction (subject vs. object) and, in addition, how the *which* phrase was disambiguated (2). If the processor searches for antecedents via a cue-based retrieval mechanism with limited focus of attention (e.g., Lewis & Vasishth, 2005), then pronominal continuations with limited cue information (*i. one/ ii. of them* in (2)) of *which* phrases should elicit greater processing costs than continuations that specify a complete NP (*iii. guest in (2)*), when OB and AD conflict. In both experiments, residual reading times revealed a penalty when the structurally preferred location for antecedents lacked overt alternatives for both pronominal continuations over NP counterparts, but that cue-rich plural pronouns (*of them*) elicited an earlier and sustained effect (regions 7–8) compared to the effect for ambiguous pronouns (*one*), which only appeared in region 8. However, a complete NP (*guest*) elicited no penalty, consistent with content-addressable retrieval. In sum, when the processor uses cues to retrieve an antecedent, it prefers one from object position (OB), and is facilitated when the antecedent is discourse-new (AD). A final self-paced reading study (N=28) tested whether the delay for retrieving indefinite antecedents (*guest*) from cue-poor probes (*one*) would be facilitated by contexts implicating multiple salient entities (e.g., a party scenario) over neutral control contexts. The delayed advantage for object disjunctions was replicated in the final region. However, even though the efficacy of these contexts was confirmed in an exit study, supporting contexts showed no facilitation for retrieval of indefinite antecedents.

These findings support a cue-based, content-addressable model of retrieval for D-linked phrases in which the processor (i) relies on a structural constraint (OB) to guide the focus of attention, (ii) avoids generating discourse alternatives during retrieval, if possible, and (iii) delays resolution of cue-poor probes, regardless of supportive context.

(1) a. A guest talked to Bill or Sue, but I don't remember which / which one.
      b. Bill or Sue talked to a guest, but I don't remember which / which one.
 (2) a. |1 A guest |2 talked to |3 Bill or Sue, |4 but I don't |5 remember which …
       b. |1 Bill or Sue |2 talked to |3 a guest, |4 but I don't |5 remember which …
          |6 (i.) one / (ii.) of them / (iii.) guest |7 it was |8 in the end.

New evidence on D-linking
Grant Goodall (University of California, San Diego)
ggoodall@ucsd.edu

It is well known that extraction out of certain islands is more acceptable with "D-linked" *wh*-phrases (e.g. *which of the books*) than with bare *wh*-words (e.g. *what*). There are three main families of explanations for this: **Semantic** accounts (Szabolcsi & Zwarts 1993, etc.) claim that D-linking increases the granularity of the extractee, which facilitates the Boolean operation required by the island-inducing operator. **Syntactic** accounts (Rizzi 2002, etc.) claim that D-linked *wh*-phrases may be interpreted as topics and thus may bind their gaps without violating the island. **Processing** accounts (Hofmeister 2007, etc.) claim that D-linked *wh*-phrases are more easily retained in working memory and are thus more easily reintegrated at the point of the gap.

The processing account alone predicts a D-linking effect in both island and non-island environments. If integration of a D-linked filler at the gap site is easier for processing, this effect should in principle be detectable in any environment. In the other two accounts, no advantage for D-linked fillers is predicted outside of islands. Here we test this prediction by means of an acceptability experiment. Previous attempts along these lines have either been inconclusive or failed to find a D-linking effect in non-islands. The danger of a ceiling effect with non-islands is very real, however, so we attempt to avoid this by using a fully counterbalanced design and many filler items of varying acceptability, including many of very high acceptability.

Method/Materials. 48 participants judged acceptability of sentences with a 7-point scale in a laboratory setting. Materials (samples in (1) were *wh*-questions with extraction out of complex NPs (a CNPC island), embedded *wh*-questions (a *wh*-island), and *that*-clauses (a non-island), with either bare or D-linked fillers.

(1)a. **What / Which of the cars** do you believe the claim that he might buy?
   b. **What / Which of the cars** do you wonder who might buy?
   c. **What / Which of the cars** do you believe that he might buy?

12 counterbalanced and pseudo-randomized lists were created with a Latin square design and 4 tokens of each condition, plus 81 fillers (105 total items). An additional 12 lists were created by reversing the order of items. Two participants were randomly assigned to each of the 24 lists.
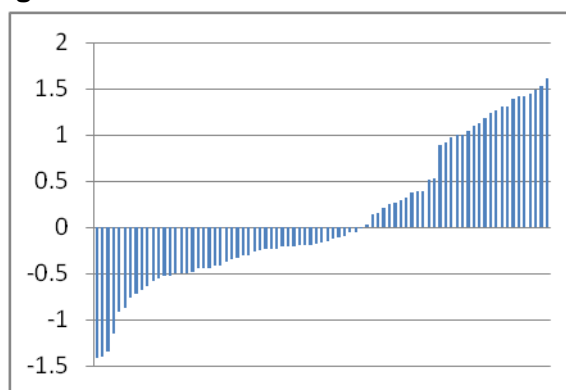
Figure 1. Z-score means (with SE)
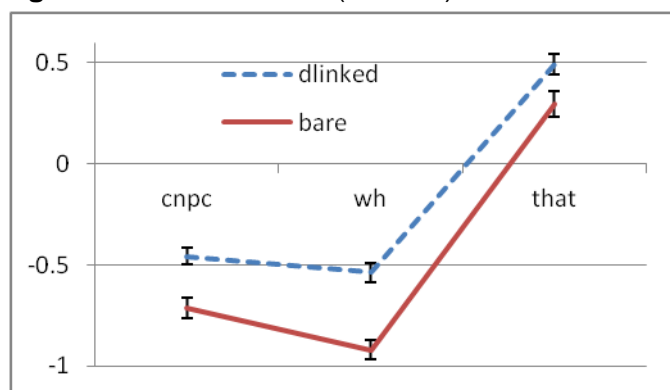
Figure 2. Distribution of z-score means of fillers

Results. Results (see Fig. 1) show main effects for filler-type (p< .001) and structure-type (p<.001), but no interaction between the two (p=.134). Post-hoc comparison shows a significant difference between bare and D-linked fillers in all three structures: CNPC (p<.001), *wh*-island (p<.001), and *that*-clause (p=.019). **The prediction made by processing accounts of D-linking thus appears to be correct.** A comparison of Fig. 1 and Fig. 2 reveals that the means for the *that*-clause cases are within the mid-range of the fillers. This may have helped avoid a ceiling effect and allowed the D-linking effect to become evident in *that*-clauses.

Poster Session 2

Hidden factors in the production of grammaticality judgments

Gisbert Fanselow, Jana Häussler (University of Postdam), & Thomas Weskott (University of Göttingen)
jana.haeussler@uni-potsdam.de

Grammaticality judgments are the key tool in syntactic research, whether obtained informally or by way of controlled experiments. Still, we know little about the processes and factors involved in making a grammaticality judgment. Since Schuetze (1996) there is a growing interest in this issue. The present study contributes to this body of work. We ran four questionnaire studies. Three of them combine a binary grammaticality judgment and an additional judgment on a 4pt scale. The second judgment required the participants to judge how certain they feel about their grammaticality judgment (Qu1), how easy/difficult they consider the sentence to process (Qu2), and how many people they think might share their judgment (Qu3). The last questionnaire (Qu4) examined the same materials without a secondary judgment. The materials comprise four subexperiments examining adverb scope (ADVSC), scrambling (SCR), split-NPs (SNP), and superiority violations (SUV), and six ungrammatical pure filler items.

Overall, participants tend to be more tolerant when allowed to express doubt my means of an extra-judgment. However, details differ depending on the type of secondary judgment and the type of construction. A mixed-effects logistic regression including participants and items as random (intercept) effects attests a main effect of secondary judgment for Qu1 and Qu2 but not for Qu3 (cf. Table 1). Correlations between the two judgments also vary across type of secondary judgment and subexperiment (Kendall's tau, cf. Table 2).

Table 1. Fixed effects in LME (LogLik = -13505)

	Estimate	SE	z	p
Intercept	.15	.09	1.65	.10
Qu1 (uncertainty) vs. Qu4 (single rating)	.32	.13	1.43	.02*
Qu2 (perceived difficulty) vs. Qu4 (single rating)	.32	.13	1.49	.01*
Qu3 (consensus estimate) vs. Qu4 (single rating)	.03	.13	.23	.82

Table 2. Rank correlations for the two judgments in each questionnaire and each subexperiment

	ADVSC	SCR	SNP	SUV
Qu1: acceptance rate and uncertainty	-.09*	-.40*	-.08+	-.20*
Qu2: acceptance rate and perceived difficulty	-.24*	-.33*	-.28*	-.50*
Qu3: acceptance rate and consensus estimate	-.07	.44*	-.08	-.15*

The possibility to give a certainty judgment increased acceptance rates in all experiments except for constructions that were overwhelmingly rejected, e.g. split-NPs with a number conflict (*Professoren*PL *kennt sie keinen*SG `As for professors, she knows none'). Grammaticality judgments and certainty judgments `correlate though not perfectly. Closer inspection reveals that high acceptance rates and low acceptance rates come along with high certainty whereas uncertainty increases slightly for acceptance rates in the middle range. Rating the difficulty of a sentence had a different effect. Again, acceptance rates are higher than in the single-judgment questionnaire; but this time, the increase affects also ungrammatical sentences (e.g. split-NPs mentioned above) but spares sentences containing a superiority violation. Grammaticality judgments and difficulty ratings correlate in all subexperiments. Finally, estimates about the number of people who share the judgment hardly affected the grammaticality judgment. Grammaticality judgments and consensus estimates do not correlate—except for the scrambling subexperiment. Lower acceptability goes along with a lower consensus estimate. Participants rejected sentences with object-before-subject order more often and they estimated that many people share this judgment. To our knowledge, our experiments are the first to systematically investigate the influence of secondary task on acceptability judgments. Taken together they show that hidden factors can be at least in part captured by means of an explicit rating task.

The underlying cognitive components of sentence processing: Not all P600s are alike
Polly O'Rourke (University of Maryland Center for the Advanced Study of Language)
porourke@casl.umd.edu

Research using the event-related potential (ERP) technique has shown that the P600, a positive component maximal between 500 and 800 ms over posterior scalp sites, is elicited grammatical violations of all kinds but also by well-formed sentences which present increased difficulty during parsing. This effect is elicited by the resolution of syntactic ambiguity (i.e. the disambiguating word in a garden-path sentence) (Gouvea, Phillips, Kazanina & Poeppel. 2010, Osterhout, Holcomb & Swinney 1994), as well as the completion of filler-gap dependencies (Gouvea et al. 2010, Kaan, Harris, Gibson & Holcomb 2000). The P600 effect is typically larger, in terms of distribution and duration, for garden-path sentences than for object relative structures that contain no temporary ambiguity (Gouvea et al. 2010). The current study explored the possibility that this difference does not merely reflect a difference in effect magnitude but rather a qualitative difference reflecting distinct cognitive substrates for the processing temporary ambiguity (which requires reanalysis) and syntactic complexity (which places increased demands on the parser but does not require reanalysis). The specific question was whether the P600 effects for the two structure types would correlate with different measures of working memory. Working memory (WM) is "a multicomponent system responsible for active maintenance of information in the face of ongoing processing and/or distraction" (Conway, Kane, Bunting, Hambrick, Wilhelm & Engle 2005, p. 770). Four measures were included which indexed different WM mechanisms: reading span (storage and sentence processing under divided attention), operation span (storage and non-linguistic processing under divided attention), *n*-back (maintenance and regular updating of information) and anti-saccade (suppression of pre-potent response). It is known that performance on the reading span task predicts P600 effect size for garden-path sentences but not for unambiguous object relatives (Friederici, Steinhauer, Mecklinger & Meyer, 1998).

EEG was recorded while participants read garden path, object relative and control sentences (see 1-3 below) and answered comprehension questions. Most of the sentence stimuli came from that of Gouvea et al. (2010). After the EEG session, participants completed the working memory assessments. Data from 51 participants is included in this analysis.

1) The patient met the doctor and the nurse with the white dress <u>showed</u> the chart during the meeting. (Garden-Path)
2) The patient met the doctor to whom the nurse with the white dress <u>showed</u> the chart during the meeting. (Object Relative)
3) The patient met the doctor while the nurse with the white dress <u>showed</u> the chart during the meeting. (Control)

The P600s effects replicated those of Gouvea et al. (2010) such that the effects for garden-path sentences were significant in the 500-700 ms and 700-900 ms time windows while effects for object relatives were limited to the latter time frame. The effects of garden-paths were also more widely distributed. Examination of the WM tasks showed that though the two complex span tasks (reading and operation span) correlated with each other ($r = .353$, $p < .05$), neither correlated with *n*-back (a finding consistent with Kane, Conway, Miura & Colflesh, 2007). With respect to the P600, the effect for garden-paths (in the 700-900 ms window at PZ) correlated positively with reading span score ($r = .293$, $p < .05$) while the effect for object relatives correlated negatively with *n*-back accuracy ($r = -.376$, $p < .01$). There were no significant correlations involving operation span or anti-saccade. The correlations with very distinct measures of very distinct WM mechanisms (and in different directions) suggest that the difference in time course and distribution reflects a qualitative difference and provide evidence for the hypothesis that the reanalysis required by temporary syntactic ambiguity and the parsing of complex syntax elicit distinct late positive components. These results also suggest a more complex relationship between cognitive control and sentence processing than has previously been considered.

Poster Session 2

The processing of raising and nominal control
Patrick Sturt (University of Edinburgh) & Nayoung Kwon (Konkuk University)
patrick.sturt@ed.ac.uk

Almost any model of human sentence processing must assume that comprehension involves the computation of linguistic dependencies. However, it is only relatively recently that researchers have begun to turn their attention to the question of *how* linguistic dependency formation takes place, for example, what type of memory access procedure is involved [A,B], or how the target of the dependency is represented. According to one recent proposal, aspects of dependency processing, such as memory retrieval, may differ according to the type of dependency involved [C]. In the current study, we further investigate this issue by examining the processing of *indirect* dependencies, as illustrated in 1:

1a. I was touched at Liz's vow [to Kevin] to be true to himself/herself sincerely and with honesty.

1b. It was touching that Liz seemed [to Kevin] to be true to himself/herself sincerely and with honesty.

In both 1a and 1b, the reflexive *himself/herself* forms a direct dependency (call it D1) with the phonologically unexpressed subject of *to be true,* which, in turn, forms another direct dependency (D2) with the overt antecedent *Liz.* Thus, in both cases, the reflexive forms an indirect dependency with *Liz.* Examples 1a and 1b differ in that the dependency between the reflexive and *Liz* in 1a is mediated by a nominal control relation (based on the lexical properties of the noun *vow*)*,* while in 1b, it is mediated by a raising relation, licensed by the verb *seemed.* Note that the properties of D1 are identical between 1a and 1b. In particular, the overt lexical item that triggers the retrieval (i.e. the reflexive) and its retrieval cues (e.g. animacy, gender, etc) are identical, so any processing differences between 1a and 1b must be attributable to differences in representation and/or processing associated with D2.

We report two eye-tracking experiments, each with 32 subjects and 40 items. Experiment 1 examined sentences like 1a,b, crossing gender matching (match vs. mismatch of reflexive with antecedent) with structure (raising vs. control). Results showed a gender mismatch cost for both 1a and 1b, but the onset this processing difficulty (relative to initial fixation on the reflexive) differed according to structure type. The earliest indication of processing difficulty came in regression path time at the reflexive region in 1b (raising), while 1a (nominal control) showed no mismatch effect in this measure (matching x structure, p's < .05). In later processing regions, significant mismatch effects were found for both structure types.

In Experiment 2, the stimuli in 1a,b were adapted by removing the grammatically illicit antecedent (i.e. [*to Kevin* in 1a,b, was deleted). In contrast to Experiment 1, results showed a very early onset of gender mismatch difficulty for both structure types (main effect of gender-match at reflexive: p's < .05 first-fixation, first-pass, regression path on reflexive), with no indication of an interaction with structure type (F's < 1). This suggests that the onset difference found in Experiment 1 between 1a and 1b was due to the differential susceptibility to interference from the illicit antecedent in the two dependency types. We argue that the representation of the antecedent for nominal control dependencies is less determinate than that of raising dependencies. This may be due to the reliance of nominal control dependencies on lexical and semantic information (e.g. control properties of *vow* in 1a), in contrast to the more determinate nature of the raising dependency, where the target of the dependency is guaranteed to be the subject of the raising verb. The results thus provide further evidence for representational differences among dependency types, and insights into how such differences play out in on-line processing.

References
A. McElree, B., Foraker, S. & Dyer, L. (2003). *Journal of Memory and Language,* 48, 67-91
B. Vasishth, S. Bruessow, R. L. Lewis, and H. Drenhaus. *Cognitive Science*, 32(4), 2008
C. Phillips, C., Wagers, M. W., & Lau, E. F. (2011). Syntax & Semantics, vol. 37, pp. 153-186.

Biases in resolving *wh*-dependencies in a hybrid language

Dustin A. Chacón (University of Maryland), Colin Phillips (University of Maryland)
dchacon@umd.edu

Contrary to the well-established bias for short-distance resolution of filler-gap dependencies in English-type languages, Japanese speakers prefer long-distance resolution into an embedded clause, even when a local resolution is possible (Aoshima et al 2004, Omaki 2010, Nakano et al. 2002). This contrast has been taken as evidence for a universal bias to associate fronted fillers with the first verb in a sentence, since in head final languages, embedded clause verbs occur before main clause verbs. Moreover, that the effect of word order outweighs the effects of structural preferences or verb biases. However, crosslinguistic comparison introduces a number of possible differences that might explain this contrast. The semantics of the embedding verb, the type of filler-gap dependency (wh-fronting vs. scrambling), or other structural factors may influence the filler-gap resolution. We present evidence from Bangla, an Indo-Aryan language that permits a more controlled within-language comparison by admitting both verb-complement and complement-verb word orders. Results from a Question After Story task and judgment task show that verb-complement ordering reliably affects filler-gap dependency preferences. However, the strength of the effect is less categorical than in prior studies, suggesting that the unique contribution of verb-complement order may have been overstated.

Bangla is a "hybrid language", permitting either verb-complement or complement-verb orders and features obligatory *wh*-movement to post-subject position, shown in (1) and (2) (Bayer 1999, Simpson & Bhattacharya 2003). Both (1-2) are ambiguous in that the *wh*-dependency may resolve in the main or embedded clause.

(1) *Raj kɔkhon* ___ *bollo* [*je Sita* ___ *am khelo*]
    Raj when   said   that Sita   mango ate
(2) *Raj kɔkhon* ___   [*Sita* ___   *am khelo bole*] *bollo*
    Raj when   Sita   mango ate   that said
    `When did Raj say ___ that Sita ate a mango ___?'

If the verb-complement word order uniquely controls preferences for filler-gap dependency interpretation, then we predict that in (1) a Bangla speaker should prefer to resolve the filler-gap dependency in the main clause, and in (2) in the embedded clause. This is because the first verb encountered by the parser in (1) is *bollo* 'said', and in (2) *khelo* 'ate'. We devised a Question After Story task (de Villiers et al. 1990) to test this prediction. In this task, speakers watched short videos, in which a character performs an action in one location, and reports having performed the action elsewhere. Afterwards, an ambiguous *wh*-question was asked, using either of the word orders in (1-2). Answers revealed how the participant interpreted the wh-dependency. Results (16 participants, 8 stories) showed a reliable within-language effect of word order on interpretation (t(48), p = 0.04), but it was much weaker than the near categorical contrast that Omaki (2010) found in a between-language comparison of English and Japanese.

One possibility for this is that the *Comp-Verb* word order is marked in Bangla, and thus there may be difficulty for the parser in constructing the embedded clause that impacts filler-gap resolution. A judgment study investigating word order preferences confirmed this asymmetry. We found a marginally significant effect of word order, corresponding with the hypothesis that *Comp-Verb* word order is more marked (F(1, 12) = 3.19, p = 0.07). Thus, one plausible explanation for the filler-gap resolution biases in the *Comp-Verb* word order is that the Bangla parser struggles with the less frequent, marked word order, causing erratic filler-gap resolutions.

These results show that there is an impact of word order on *wh*-dependency resolution, however this preference does not uniquely determine resolution. This is supported by our results, which show that there is not a categorical distinction between the two word orders in a language-internal comparison, as found in Omaki's (2010) cross-linguistic studies. Thus, although word order is one factor in determining *wh*-dependency resolution, it does not necessarily override competing factors universally.

Poster Session 2

Argument-structure driven parsing in Tagalog

Michael Frazier, Masaya Yoshida (Northwestern University)
fraze@u.northwestern.edu

Previous literature on sentence processing has identified several structure-building biases. One is a preference toward satisfying a verb's argument structure as soon as possible, the argument attachment preference (AA, [1,2,3]). Another is a preference to attach an incoming phrase at the right edge of the existing structure, the local attachment preference (LA, [4,5]). This study provides support for the importance of argument structure in a novel domain: Tagalog argument/possessor ambiguity. Because Tagalog is a V-initial language, argument structure information is available from the beginning of a sentence. In the Recent Perfective (RP) form, the case morphology on the arguments of some verbs is neutralized, surfacing as the genitive (GEN) form. Thus, when a sentence contains sequential genitive NPs, it is ambiguous whether these genitive NPs are arguments of the verb or one of the NPs is a possessive modifier of the other NP. In such sentences the parser must determine the grammatical function and thematic roles of the NPs without the help of case morphology. When a transitive verb like choose in the RP form is followed by three genitive NPs, the grammar does not determine whether the structure is (1) [V [NP1 NP2] NP3] or (2) [V NP1 [NP2 NP3]], as both are grammatical in Tagalog. Here, the two parsing biases give different predictions: AA predicts that the parser, by default, assigns incoming NPs to θ-positions and thus NP1 and NP2 will be initially parsed as arguments of the verb. Therefore if NP2 turns out to be a possessive modifier of NP1 as in (2), a garden path effect will occur. Thus if NP3 is not a plausible possessor for NP2, NP3 should cause reading time slowdown. In contrast, LA predicts that the NP2 is most preferentially attached to NP1 as NP1 is the most local attachment site of NP2. Thus, the garden path effect will arise when the gen-NP2 is not a plausible possessor for the gen-NP1, giving rise to the slow reading time of NP2.. In a word-by-word self-paced reading experiment, 38 Tagalog speakers in the Philippines read 32 sentences in which three NPs followed the verb. Case Form of NPs (Neutralized vs. Non-Neutralized), Plausibility (Possessed-Possessor sequence is plausible vs. implausible) and position of complex NP (first vs. second argument after the verb) are manipulated as independent factors.

(3) V   NP$_1$-GEN   [NP$_2$-NOM   NP$_3$-GEN]   (a/b: complex right NP plausible/implausible)
     V   [NP$_1$-GEN   NP$_2$-GEN]   NP$_3$-NOM   (c/d: complex left NP plausible/implausible)
     V-**RP** NP$_1$-GEN   NP$_2$-GEN   NP$_3$-GEN   (e/f: ambiguous; plausible complex NP left/right)

In the normal perfective, 3-NP sequences are unambiguous if the nominative-marked NP follows the genitive-marked NP. So conditions (3a-d) have only a single parse. In (3a/b), the complex NP comes second, while in (3c/d), it come first. Plausibility (Possessed-Possessor sequence is plausible vs. implausible) was also varied in the non-case-neutralizing conditions, such that in (3a/d), the complex NP (possessed-possessor) is plausible, while in (3b/c) it is implausible. Due to the absence of case cues, plausibility is indistinguishable from position of the complex NP for the RP conditions (3e/f). In RP conditions, then, NP2 can be parsed either as the possessor of NP1 as in (1) or an argument of the verb and the possessed of NP3 as in (2), with Plausibility serving as a probe to determine which of the two possible structures the parser builds by default. In (e), parse (1) is more plausible, while in (f), parse (2) is.

RTs at NP3 are significantly slower in the RP when NP3 is not a plausible possessor for NP2 (p<.05; linear mixed model, t=-3.11, df=7). No significant effects are found at NP2 in any condition. This suggests that the parser's AA bias is stronger than its LA bias: the parser preferentially attaches case-ambiguous NPs to argument slots rather than the structurally most local slots. This suggests an important role for the lexical information of the verb in biasing parsing decisions. Complementarily, the presence of this effect only in RP conditions, where the verbal morphology signals that case-neutralization to GEN will occur, is evidence for the parser's sensitivity to grammatically encoded argument-structure cues in real time.

References [1] Pritchett (1988). Language 64. [2] Speer & Clifton (1998). Memory & Cognition 26(5). [3] Gilboy et al. (1995). Cognition 54:131-167. [4] Kimball (1973). Cognition 2:15–47. [5] Phillips & Gibson (1997). JPR 26.3.

Effects of syntactic complexity and animacy on the initiation times for head-final relative clauses

Chien-Jer Charles Lin (Indiana University)
chiclin@indiana.edu

Introduction. While a uniform processing advantage of subject relatives (SRs) over object relatives (ORs) has been observed both in the comprehension and the production of head-initial relative clauses (RCs), a mixture of SR and OR advantages has been reported for the comprehension of head-final RCs. Factors argued to be responsible for such processing asymmetries include structural frequency [1], temporary ambiguity [2], phrasal complexity [3], linear locality [4], animacy of the NPs [5,6], and canonicity of word orders [7], among others. Most studies thus far have focused on the *comprehension* of head-final RCs, where temporary structural ambiguity exists before the head NP. To get around the issue of parsing uncertainty, and to examine the correspondences between RC comprehension and production, the current study investigated initiation times for producing head-final RCs in Mandarin Chinese. We test the hypothesis that SRs are syntactically more complex than ORs as the gap in an SR is outside of a VP but that in an OR is internal to the VP. Phrasal integrity predicts that SRs be easier to produce than ORs (see [3, 8] for similar accounts for an SR comprehension advantage in Korean & Japanese). Both linear locality [4] and dominant word order [7] accounts predict ORs be easier to produce in Mandarin Chinese.

Design. Corpus distributions [5] show that SRs (61%) outnumber ORs (39%) at both matrix subject and object positions. Crucially, when the RC has an inanimate head NP and an animate embedded NP (e.g., SR: *the stone that tripped the girl*; OR: *the stone that the girl picked*), ORs (75%) outnumber SRs (25%). Two RC production experiments were conducted, in which participants (N=42) were prompted to produce an RC such as (2-3) based on two given sentences and a prompt in (1). Exp. 1 manipulated extraction types (SR/OR) and matrix positions (subject vs. object modification) of the RCs. Exp. 2 manipulated extraction types (SR/OR) and the animacy of the NPs in the RCs. Initiation times [9] and the produced RCs were recorded.

1.   A student assisted a professor. The student$_{(SR)}$/professor$_{(OR)}$ visited a famous dancer.
       Prompt: _____ *visited a famous dancer.*
2.   Target SR: [$_{CP}$ _ [$_{VP}$ assist professor]] rel student$_{head}$ 'the student who assisted the professor'
3.   Target OR: [$_{CP}$ student [$_{VP}$ assist _]] rel professor$_{head}$ 'the professor who the student assisted'

Results. Consistent with the corpus distributions, Exp. 1 found shorter initiation times for SRs in both matrix positions (p=.007). Significant interactions indicated greater differences at the matrix subject position (p=.034). Exp. 2 found shorter initiation times for SRs than ORs (p=.053). Animate head NPs require shorter planning times than inanimate head NPs (p=.03). Crucially, initiation times for SRs were still shorter than ORs when the RC had an inanimate head NP and an animate embedded NP (p=.006), showing an SR advantage that is not predicted by corpus distributions. (All statistics were based on ANOVAs with repeated measures.)

Discussion. Since our RC production tasks already provided the lexical items, their grammatical functions, and the target RC types, the initiation times observed are most sensitive to constituent assembly [10], syntactic complexity, and plausibility. Our results suggested that SRs are processed with greater ease because less complicated syntactic structures (e.g., in terms of phrasal integrity) are found in SRs than in ORs. This effect is independent of frequency and plausibility information based on corpus distributions. It also could not be accounted for by effects of dominant word orders [7] and linear locality [4].

References
[1] Chen et al., 2012. [2] Lin & Bever, 2011. [3] Kwon et al., 2010, Language. [4] Gibson & H. Wu, 2013, LCP. [5] F. Wu et al., 2012, LCP. [6] Hsiao & MacDonald, 2011, CUNY poster. [7] Su et al., 2007. Bran & Language. [8] Ueno & Garnsey, 2008, LCP. [9] Ferreira, 1991, JML. [10] Bock & Levelt, 1994. Handbook of Psycholinguistics.

Collectivity and concreteness in optional Persian number agreement
Aazam Feizmohammadpour and Wind Cowles (University of Florida)
afeiz@ufl.edu

Like many languages, Persian exhibits subject-verb number agreement. However, this agreement in Persian becomes optional when the subject is an inanimate plural noun. This allows us to investigate the factors influencing subject-verb agreement without needing to elicit ungrammatical structures from speakers. In these two experiments, factors previously found to influence agreement error rates (collectivity and concreteness) were tested in Persian.

Experiment 1:

The goal of this experiment was to test for effects of collectivity on subject-verb agreement (e.g. Humphreys & Bock, 2005) when that agreement is optional. 24 native speakers of Persian were provided a series of preambles and instructed to repeat the preambles aloud and then complete the sentences. The preambles consisted of an inanimate plural subject, a prepositional phrase modifier and either a past participle or an adjective, creating two similar structures, passive and copular. The prepositions were manipulated so that the prepositional phrases modifying the subject head noun emphasized the unity (e.g. *"the cups **near** the saucers"*) or individuality (e.g. *"the cups **on** the saucers"*) of the entities in subject position.

Responses were coded for singular *vs.* plural morphology on the auxiliary verb (e.g. *shod vs. shod-æn*). The results show that for both structure types, singular verbs were used more often when the plural subject was perceived as a unit (Unity: 43% singular verbs) rather than as a collection of individuals (Individuality: 29% singular verbs), showing that the perception of the subject as more or less plural influences grammatical number agreement in Persian.

Experiment 2:

The goal of this experiment was to test for effects of concreteness vs. abstractness of the subject noun (e.g. Eberhard, 1999) in optional subject-verb agreement. 24 native speakers of Persian were given a series of preambles and instructed to repeat the preambles aloud and then complete the sentences. In critical trials, the preambles consisted of an inanimate plural subject noun (whose referent was either concrete or abstract) and a past participle or adjective.
Responses were again coded for singular *vs.* plural morphology on the auxiliary verb, and the results show more singular verb use with abstract subjects (67%) compared to concrete ones (46%) for both copular and passive sentences.

Taken together, these results confirm the influence of conceptual factors on subject-verb agreement, showing in fully grammatical contexts, agreement morphology is subject to the perceived unity and abstractness of the subject.

E1 Example: Subject {Individuality/Unit} Preambles (with possible participant responses)
negin-ha-ye   {ruye/kenare} ængoshtær-ha   pæsændid-e (shod   / shod-æn)
gem-PL-ez   {on/near}   ring-PL   chose-pp   (became.3SG / became-3PL)
'the gems on/near the rings (was/were) chosen'

E2 Example: Subject {Abstract/Concrete} Preambles (with possible participant responses)
{omid-ha / kamputer-ha}   dad-e   (shod   / shod-æn)
{hope-PL / computer-PL}   gave-PP   (became-3SG / became-3PL)
'the hopes / computers (was/were) given'

Number agreement without surface syntax

Ming Xiang, Genna Vegh (University of Chicago)
mxiang@uchicago.edu

Although number-mismatch in overt syntax is ungrammatical in English ("*John is teachers*"), VP ellipsis appears insensitive to such mismatch. Sentences like "*They are teachers, and John is too*" are acceptable, although the predicate of the ellipsis site must be recovered from the antecedent ("…*and John is* ~~*teachers*~~ *too*"), which presents a number-mismatched candidate. In an ERP experiment, the current study showed that such number-mismatch, though grammatical, is nevertheless costly in online processing, suggesting that number agreement, and therefore syntactic computations, are being carried out at the ellipsis site.

Two general classes of proposals could potentially derive grammatical number-mismatch for VP ellipsis. The first assumes that number agreement is calculated in ellipsis, but when the number feature on the antecedent mismatches with that of the ellipsis site, additional operations are evoked to fix the mismatch (Merchant, 2011). This approach could predict extra processing cost for the ellipsis with a number-mismatched antecedent. The second class of proposals, on the other hand, denies any actual mismatch, either by assuming agreement is only calculated in morpho-phonology, not in syntax per se (Bobaljik, 2008) or ellipsis is resolved purely in semantics/discourse, not in syntax (Hardt 1993). This would predict no processing cost for the seeming number-mismatch in ellipsis, since there is no mismatch to start with.

160 sets of 4-condition items were constructed, based on a 2x2 design: the number feature of the subject at the ellipsis site is either singular or plural); and the number feature in the antecedent either matches or mismatches with the subject at the ellipsis site.

(1)a. The worker here is a writer. The worker there is too. (Sg. subject, Match)
   b. The workers here are writers. The worker there is too.(Sg. subject, Mismatch)
   c. The workers here are writers. The workers there are too.(Pl. subject, Match)
   d. The worker here is a writer. The workers there are too.(Pl. subject, Mismatch)

Behavioral acceptability norming (n=20) showed no difference between conditions (ps>.1; ranging between 5.7-6 on a 7-point acceptability scale). For ERP testing, 160 bi-clausal controls were created that contain similar structures, but no ellipsis. Since "too" is the first word in the second clause that unambiguously signals an ellipsis structure, we time-locked ERPs to the onset of this word. Preliminary results (n=20) showed that there is a main effect of Number feature during the N400 time window (350-450ms) at the central-frontal and posterior scalp sites (ps<.01), such that the ellipsis site with a plural subject elicited a larger N400 relative to that with a singular subject. This holds for the comparison between the two Match conditions (1c vs. 1a, ps<.001), and also between the two Mismatched conditions (1d vs. 1b, ps<.05). For the effect of Match, when the ellipsis has a singular subject (1a & 1b), the mismatched condition (1b) elicited a larger N400 between 350-400ms, at central-frontal and posterior sites (ps<.01), and no other difference was observed. When the ellipsis has a plural subject (1c & 1d), there is a small central-frontal N400 to the mismatched (1d), but only approaching significance at central-frontal site (p=.1); however, (1d) also elicited an extended positivity between 600-1000ms at the right anterior region (p<.001).

These results show that number agreement without surface syntax is being computed in online processing, suggesting that syntactic representations are accessed in ellipsis resolution. Furthermore, such computation makes distinctions between plural and singular features: (i) plural agreement at the ellipsis site is in general more costly compared to singulars; and (ii) mismatched plural agreement (compared to the matched plural one) elicited different ERP patterns from mismatched singular agreement (compared to the matched singular one). Unlike in the current results, number agreement in previous ERP studies have generally elicited LAN and P600 effect. A critical distinction to consider is that previous studies have all looked at agreement in overt syntax; therefore, the ERP components observed before could be largely triggered by morpho-syntactic processing of the overt inflectional morphemes, which doesn't play any role in the current study.

Poster Session 2

Predictability effects of case-marked direct objects: Evidence from Romanian

Sofiana Chiriacescu (University of Köln, Transilvania University of Brasov)
sofiana.chiriacescu@uni-koeln.de

Introduction. Language processing requires both the integration of new information and the anticipation of what is likely to be discussed next (Levy, 2008). Linguistic and psycholinguistic studies showed that several factors, such as prominent syntactic positions, different verb types, or coherence relations influence the *predictability* of upcoming information (Grosz, Aravind & Weinstein 1995; Kehler, Kertz, Rohde & Elman 2008). In this paper we report the results of a multi-sentence story-continuation study on Romanian, which shows that *case-marking on direct objects* influences higher-level discourse expectations.

When referring to human entities, sentential direct objects in **Romanian** are either preceded by the accusative case-marker (ACC, *pe*), or they remain zero-marked (e.g. *pe a child* vs. *a child* in ex.1). Animacy, referentiality and specificity are considered the main triggering factors of accusative case-marking in Romanian and cross-linguistically (Aissen 2003). However, these factors cannot explain such cases in which the case-marker is optional in Romanian (ex.1). **The present study** investigates whether referents introduced by the case-marker and referents introduced by ZERO (i.e. the simple definite or indefinite article) influence cross-sentence discourse expectations, such as (i) frequency of re-mention, and (ii) likelihood of becoming the upcoming topic. **Design.** We used a story-continuation task in which participants (n=24) read two-sentence stories and provided five logical and natural-sounding sentence continuations to each of the stories.

(1) Sample experimental item (English translation)

ACC-condition	ZERO-condition
Yesterday evening was so warm that **Graur$_1$** decided to hang out with friends at the local coffee shop. On his way downtown, **he$_1$** saw <u>**ACC-a child**</u>$_2$ coming down the street.	Yesterday evening was so warm that **Graur$_1$** decided to hang out with friends at the local coffee shop. On his way downtown, **he$_1$** saw <u>**a child**</u>$_2$ coming down the street.

The first sentence set the context and introduced a *global topic* (e.g. *Graur$_1$*, mentioned at least twice in each experimental item). The second sentence introduced the *ACC-marked* or *ZERO-marked* target referent in a *non-prominent position (e.g. a child$_2$*, mentioned once as a direct object). The verbs used in the study were neutral with respect to IC bias (McKoon, Greene & Ratcliff, 1993). **The results** show that case-marked referents were picked up in the subsequent discourse more often than their zero-marked counterparts (in 90% vs.10% of the cases), and were also more prone to become the upcoming topic (in 85% out of 100% of responses). However, case-marking did not influence the anaphoric expression used to pick up for the first time the target referents (i.e. definite noun phrases were used in both conditions).

Conclusions. First, the results of the experiment indicate that effects of predictability in terms of frequency of re-mention and upcoming topic shift do exist, although they do not appear to influence referring-expression choice. Second, contrary to previous investigations on differential object marking in Romanian (Aissen, 2003), this phenomenon is not (only) a sentence-level phenomenon. We argue that, at least in Romanian, any account on accusative case-marking has to integrate the observation that the marker functions as a discourse structuring device as well. We expect that case-marking in other languages fulfills a similar function.

References. [1] Aissen, J. (2003). Differential object marking: Iconicity vs. economy. [2] Grosz, Aravind & Weinstein. (1995). Centering: a framework for modeling the local coherence of discourse. *Computational Linguistics, 21*, 203-226. [3] Kehler, Kertz, Rohde & Elman. (2008). Coherence and coreference revisited. *Journal of Semantics, 25*, 1–44. [4] Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*, 1126–1177. [5] McKoon, Greene & Ratcliff. (1993). Discourse Models, Pronoun Resolution, and the Implicit Causality of Verbs. *Journal of Experimental Psychology, 19(5)*, 1040-1052.

The locus and nature of the object-extracted relative clause penalty
Jeffrey Witzel (University of Texas at Arlington) & Kenneth Forster (University of Arizona)
jeffrey.witzel@uta.edu

A well-established finding in psycholinguistics is the asymmetry in the processing of English subject- and object-extracted relative clauses (SRCs and ORCs) as in the following sentences:

> **SRC sentence:** The soldier *who roughly pushed the sailor* smashed a bottle against the bar.
> **ORC sentence:** The soldier *who the sailor roughly pushed* smashed a bottle against the bar.

Across domains of psycholinguistic inquiry, ORC sentences have been found to be more difficult to process than SRC sentences. Memory-*retrieval-based* accounts attribute this ORC penalty to costs associated with retrieving the modified NP (*The soldier*) for integration at the RC verb (*pushed*), which is done over greater distance in ORCs than in SRCs (Gordon et al., 2006; Grodner and Gibson, 2005). Syntactic-*expectancy-based* accounts, however, hold that this penalty is due to reanalysis of a default SRC parse, with this default attributed to structural frequency (Gennari & MacDonald, 2008; Levy, 2008) or to constraints on structural processing (Clifton & Frazier, 1989; Traxler et al., 2005). Finally, *hybrid* accounts posit core roles for both of these sources of processing difficulty (Lewis & Vasishht, 2005; Staub, 2010). One way to test among accounts is to examine the locus of processing difficulty in ORC sentences (Staub, 2010). Under retrieval-based accounts, this difficulty should occur primarily at the ORC verb (*pushed*), when retrieval/integration processes are engaged. Expectancy-based accounts, however, predict processing difficulty mainly at the ORC subject (*the sailor*), the element that disconfirms structure-based expectancies and triggers reanalysis. Hybrid accounts would be supported by independent processing costs at both of these points. The findings to date, however, have not revealed a consistent locus of the ORC penalty (compare e.g., Forster et al., 2009; Grodner and Gibson, 2005; Staub, 2010). In light of these conflicting findings, the present study investigated the locus of processing difficulty for ORCs under different task conditions.

Experiment 1 (*N*=36) used eye tracking to investigate SRC and ORC sentences as in the examples above. The sentences were comparable to those in Staub (2010, Experiment 1), but included an adverb (*roughly*) in the RC. In ORC sentences, this adverb created a "buffer" between the ORC subject (*the sailor*) and verb (*pushed*) to allow for clearer indications of independent processing costs at these points. The results revealed processing difficulty for ORC sentences both at the beginning of the RC and at/after the RC verb. However, there were clear differences in the strength and timing of the processing costs at these positions. At the ORC subject and immediately following adverb, there was a robust ORC penalty under multiple first-pass reading measures. By contrast, processing difficulty at the ORC verb was shown only by a relatively weak effect under a single measure – first fixation duration – with robust indications of the ORC penalty picking up again at the immediately following main clause verb (*smashed*). Experiment 2 (*N*=48) tested the same sentences using the L-maze task (Forster et al., 2009; Witzel et al., 2012). This task engages syntactic processing, but does not require careful tracking of verbs and their arguments. Under these task conditions, only ORC penalty effects related to disconfirmed syntactic expectancies were predicted. Consistent with this prediction, an ORC penalty was revealed only at the article in the ORC subject (*the* sailor).

Taken together, these findings provide clear support for hybrid models of the ORC penalty. The eye-tracking experiment revealed processing difficulty at the ORC subject and ORC verb, suggesting costs for both disconfirmed syntactic expectancies and long-distance, retrieval-based integration. The L-maze experiment further indicated that processing costs at the ORC subject and ORC verb are isolable under certain task conditions and thus that they relate to independent sources. Moreover, the fact the ORC penalty in the L-maze experiment was restricted to the article in the ORC subject – the first element that can disconfirm an SRC parse and trigger reanalysis – suggests that processing costs early in the ORC relate specifically to the violation of structural expectancies.

Who did what to whom? An investigation of syntactic reanalysis in English and Mandarin

Yi Ting Huang[1], Xiangzhi Meng[2] & Kathryn Leech[1] ([1]University of Maryland College Park, [2]Peking Univeristy)
ythuang1@umd.edu

During language comprehension, cues that are available early in utterances may lead listeners to commit to interpretations that are inconsistent with later content. For example, NP1s in passives may be initially misanalyzed as agents because of a frequency advantage for actives. Thus successful interpretation requires listeners to revise NP1s as themes following the past participle (-en). Prior cross-linguistic work has shown that linguistic cues can trigger real-time predictions of thematic roles (Kamide et al., 2003), but it remains unknown what makes certain cues informative and how predictive processes interact with syntactic reanalysis. One possibility is that more frequent cues drive reanalysis. Alternatively, it may be that more valid cues serve as effective triggers. The current study distinguishes between these possibilities by comparing interpretations of actives and passives in English and Mandarin. While passives are less frequent than actives in both languages, this disparity is 10 times greater in Mandarin than in English. In contrast, while Mandarin recruits a dedicated passive marker (BEI), those in English are often used for the past tense. Thus this cross-linguistic comparison offers a window into which cues (frequent or valid) offer a stronger basis for syntactic reanalysis.

Using a visual-world eye-tracking paradigm, English- (n = 29) and Mandarin-speakers (n = 34) were presented with sentences like (1) and (2) while their eye-movements were measured to displays featuring expressed nouns (SEAL), likely agents (SHARK), and likely themes (FISH).

(1) Expressed. a. English: The seal is quickly eat<u>en by</u> it   b. Mandarin: Seal <u>BEI</u> it quickly eat
            c. English: The seal is quickly eat<u>ing</u> it   d. Mandarin: Seal <u>BA</u> it quickly eat

(2) Pronoun.   a. English: It is quickly eat<u>en by</u> the seal   b. Mandarin: It <u>BEI</u> seal quickly eat
            c. English: It is quickly eat<u>ing</u> the seal   d. Mandarin: It <u>BA</u> seal quickly eat

Unlike in (1), the identity of NP1 in (2) is initially ambiguous ('it'). For passives, this uncertainty may prevent an early bias to construe NP1 as an agent, allowing it to be assigned as a theme later in the sentence, without reanalysis. This would lead to equivalent accuracy (i.e., correct ID of pronouns) for actives and passives in both languages. Critically, the need for reanalysis in (1) generates distinct predictions about which cues are more effective. If frequent cues have an advantage, then accuracy for passives should be higher in English than Mandarin. If, however, valid cues have an advantage, then the opposite should be true.

Analysis of eye-movements revealed that all participants used language-specific cues to incrementally assign thematic roles. Following the past participle (eat<u>en</u>), English-speakers correctly looked to agents in the expressed condition (1a) and themes in the pronoun condition (2a). Mandarin-speakers revealed parallel patterns after BEI (1b and 2b). Similarly, following the progressive suffix (eat<u>ing</u>), English-speakers looked to themes in the expressed condition (1c) and agents in the pronoun condition (2c). Mandarin-speakers revealed parallel patterns after BA (1d and 2d). This led to predicted interactions between construction (active vs. passive) and NP1 status (expressed vs. pronoun) in regions following NP2 (all p's < .01). Critically, analysis of correct actions (in bold) revealed that interpretations were also influenced by cross-linguistic differences. In the expressed condition, Mandarin-speakers were less accurate with passives than actives (p < .01), but this asymmetry was absent in English-speakers (p >.90). However, when syntactic reanalysis was unnecessary in the pronoun condition, accuracy was comparable across both constructions and populations (all p's > .50). Altogether these findings suggest that frequent cues may trump valid cues during syntactic reanalysis.

% of actions with expressed noun + other object (ID of 'it')	Expressed ('the seal is…')			Pronoun ('it is…')		
	+ AGENT	+ THEME	+ nothing	+ AGENT	+ THEME	+ nothing
English: Passive | Active	**91** | 0	4 | **91**	5 | 9	8 | **81**	**85** | 8	7 | 11
Mandarin: Passive | Active	**79** | 4	17 | **95**	4 | 1	8 | **77**	**78** | 7	14 | 16

The use of non-structural cues in reflexive resolution: Evidence from eye-tracking

Lena Benz, Lena Jäger, Shravan Vasishth (University of Potsdam) & Philip Hofmeister (University of Essex)
lenabenz@gmail.com

Whenever a reflexive pronoun in on-line sentence comprehension is encountered, a search for the reflexive's antecedent in working memory takes place. It is an open question, however, whether the initial set of possible antecedents is determined by syntactic information alone such that only nouns in a structurally appropriate (accessible) position are considered as possible antecedents or whether all preceding nouns in a sentence — even those in structurally inappropriate (inaccessible) positions — occur in the initial candidate set. Whereas [1,2] found no evidence that structurally inaccessible antecedents affect early measures of on-line processing, [3,4,5] found processing difficulties when an inaccessible antecedent matched in gender with the reflexive.
In the present study, we investigate the impact of linguistic focus, i.e., the salience of accessible and inaccessible antecedents, by manipulating the location of focus directly in the target sentences using overt focus inducers (even/only) which precede the element they are associated with. Our eye-tracking study on reflexives in English (n=34) had a 3x2 design with the factors focus (no/antecedent/intruder) and gender match (single/double).

1. {1. Only the heir who the ... | 2. The heir who only the ... | 3. The heir who the ... } {a. nephew | b. niece} envied for the heritage protected himself from being investigated on tax evasion charges.
2.
The focus was either on the accessible antecedent (1), the inaccessible antecedent (2) or there was no focus (3). The gender of the reflexive matched either with the accessible and inaccessible antecedent, (double match (1a,2a,3a) or only with the accessible antecedent (single match) (1b,2b,3b). If the parser exclusively relies on structural cues, the gender and focus manipulations on the inaccessible antecedent should not affect the reflexive's on-line resolution because it will not be considered in the antecedent search. If non-structural cues also affect the search for antecedents, we expect that the focus- and gender-marking on the inaccessible antecedent will have an effect on early measures of reflexive processing.

At the critical region containing himself, we found the following effects: (i) A focused inaccessible antecedent led to a significantly higher re-reading probability of the reflexive than a focused accessible antecedent (z= 2.8) and to a marginally higher re-reading probability than an unfocused antecedent (z=-1.8). (ii) Similarly, in total-fixation time, a focused inaccessible antecedent led to longer reading times on the reflexive compared to both, the condition with a focused accessible antecedent (t=2.6) and the condition without a focus (t=2). (iii) In the critical early measure of first-fixation duration, an interaction between match and no vs. inaccessible antecedent focus reached significance (t=-2.3): the double match condition was read faster compared to the single match condition when no element was focused; however, match had no effect when the inaccessible antecedent was focused.

These findings suggest that the parser is sensitive to non-structural cues in an early stage of reflexive resolution: We found an early speed-up at the reflexive when the accessible and inaccessible antecedents shared their gender (when no antecedents were focused). As focus eliminated this early "intrusion" effect, the results conflict with accounts, e.g. [2,5], claiming that non-structural cues are not relevant for the initial stages of reflexive resolution. In a later stage of processing, focusing inaccessible antecedents, however, proved to slow reading at the critical region, compared to conditions where it was not focused or where the true target was focused. Thus, by informationally "highlighting" a potential competitor, late stages of reflexive processing were delayed.

References. [1] Nicol & Swinney, 1989 [2] Xiang, Dillon, & Phillips, 2009 [3] Badecker & Straub, 2002 [4] Cunnings & Felser, 2011 [5] Sturt, 2003

Highs and lows in English attachment

Nino Grillo[1], Andrea Santi[2], Bruno Fernandes[1] & João Costa[1] ([1]Universidade Nova de Lisboa, [2]University College London)
nino.grillo@gmail.com

Attachment. Cuetos & Mitchell (1988) found that relative clause attachment to complex NPs differs across languages: Low Attachment (LA) is found in English (1a), and High Attachment (HA) in Spanish , among others.

The role of Pseudo Relatives. Grillo and Costa (2012) show that previous work on RC attachment overlooked the distinction between RCs and Pseudo Relatives (PR, see Cinque 1992) (2). They propose that (*everything else being equal*, e.g. prosody, frequency of Ns etc.) variation in RC attachment across languages/structures is due to asymmetric availability of PR. PRs are unavailable in LA languages (English, Romanian, German, Basque, Chinese a.o.); PRs are available in HA languages (Spanish, French, Dutch, Italian, Japanese a.o.). Grillo and Costa propose the following generalizations: i. Universally, LA preferences arise with genuine RCs and ii. When a PR is available, the parser will prefer it to an RC.

English SCs. PRs are similar to Small Clauses (SC) of the Acc-*ing* type in English (3); both denote an (undergoing) eventive argument of a predicate and have the same syntactic and semantic restrictions (e.g. progressive aspect, temporally bound by matrix T, subjects-only). PRs and SCs are distinguished only by the additional structure hosting Cs and AUXs in the case of PRs. RCs, on the other hand, are clauses that modify entities and are not bound by the same restrictions. Just like with PRs, there is an ambiguity between the SC reading and a (reduced) RC reading (4A,B) with perceptual Vs and event-introducing Ns. The ambiguity disappears with stative Vs (4C) and individual-introducing Ns (4D), which can only select for a RC. Grillo and Costa's analysis, therefore, predicts HA with complex NPs in (4A,B) type contexts and LA in (4C,D).

Experiment. In an English questionnaire we manipulated SC availability in a 2*Position*[Center Embedding, Right Branching]*x2Type*[stative, eventive] design (4). SC/RC interpretations are available in condition A and B, but only an RC reading in condition C and D. The participant was given an identical binary decision (ie, in 4 'the son runs' 'the doctor runs') across conditions. They performed the same task for 24 target sentences and 54 filler items that included garden path sentences and unambiguous full relatives. Stimuli were presented in Linger using a latin square design and full sentence presentation. As predicted by Grillo and Costa, HA preference was found in condition A (**68.0% HA**) and B (**65.2% HA**), and LA in C (**15.5% HA**) and D (**8.6% HA**). We conducted a mixed effects logistic regression using the *lmer()* function of the *lme4* package (Bates et al., 2011). *Type* and *Position* were fit as fixed factors, *subject* and *items* as random factors, and random slopes were fit for both fixed effects and their interaction. The effect of Type was significant (coefficient = 3.70, SE = 0.62, z-score = 5.999, p < .001).

Conclusions. *i.* The results strongly support the proposal that PR/SC availability is a main factor in determining Attachment and (indirectly) that the parser prefers a PR/SC reading over an RC one. Ongoing studies are testing the PR/SC interpretation preference directly and whether the present finding generalizes to full relative clauses. *ii.* Attachment is not primarily determined by syntactic position (*subject* vs. *object*) or context (*verbal* vs. *nominal*) of the complex NP *per se.* *iii.* LA emerges across the board with genuine RCs. *iv.* HA is also found in LA languages like English in contexts ambiguous between SC and RC readings.

(1)   I saw the son$_1$ of the man$_2$ that$_2$ was running
      Vi al hijo$_1$ del hombre$_2$ que$_1$ corría
(2)   Vi a [$_{PR}$ Juan que corría]
     *I saw John that run
(3)   I saw [$_{SC}$John running]

(4)   A. *John saw the son of the doctor running*
     B. *The picture of the son of the doctor running is beautiful*
     C. *John lives with the son of the doctor running*
     D. *The house of the son of the doctor running is beautiful*

Illusory NPI licensing: Now you see it, now you don't
Dan Parker, Glynis MacMillan & Colin Phillips (University of Maryland)
dparker3@umd.edu

Illusory negative polarity item (NPI) licensing has been taken as important evidence for a memory architecture for the parser that is prone to intrusion effects, in which retrieval processes are disrupted by inappropriate material [1]. This conclusion, however, has been challenged by a recent proposal which suggests that illusory NPI licensing reflects an over-extension of pragmatic inferencing mechanisms [2]. Existing evidence has come entirely from the NPI *ever*, but both accounts assume that all other NPIs should behave similarly with respect to illusory licensing. Here we provide the first direct comparison of the NPIs *ever* and *any*. Results from three studies show that whereas *ever* reliably shows the illusion, *any* resists illusory licensing. Further, these findings may support the memory retrieval account over the pragmatic account, once effects of the linear positions of the two NPIs on memory dynamics are considered. The NPIs *ever* and *any* must appear in the scope (i.e. c-command domain) of a downward- entailing operator, e.g., negation (1a). Illusory licensing manifests as eased processing or greater acceptability of illicit NPIs in the presence of a structurally inappropriate, but semantically appropriate licensor (1b). Under the retrieval account, the NPI in (1) initiates a search for a c-commanding, negative licensor. Illusions occur when the structurally inappropriate licensor is (mis-)retrieved via a partial match with negation. The pragmatic account follows the view that NPIs are not licensed by item-to-item dependencies, but rather by pragmatic inferencing [3]. Under this account, the contrastive negative inference generated from the restrictive relative clause may spuriously license the NPI.

To date, illusory licensing has only been tested for the NPI *ever*. The more prototypical NPI *any* has not been tested due to the confound of a free-choice interpretation, e.g., *Sally will marry any doctor*. Offline judgments (Experiment 1, n=18) show that the NPI interpretation of *any* can be forced when it occurs as the determiner of an abstract mass noun, e.g., *any satisfaction.* Experiment 2 (self-paced reading, n=24) directly compared the NPIs *ever* and *any* using sentences like (2)-(3), varying the presence and structural position of a potential licensor, and the NPI-type. These NPIs share identical licensing constraints in (2)-(3), and despite superficial differences in syntactic category, they are matched on the many semantic properties of NPIs addressed in detailed surveys [e.g. 4]. Results showed contrasting illusory licensing profiles for *ever* and *any*, as reflected in an interaction of illusory licensing x NPI type (LME: *pMCMC*=0.03). Pairwise comparisons showed that this interaction was driven by a significant illusory licensing effect for *ever*, (*pMCMC*=0.02), whereas no such effect was observed for *any* (*pMCMC*=0.4).

We subsequently modeled the effects with ACT-R [1,5], using identical retrieval cues for the two NPIs. Modeling revealed that the contrasting illusory licensing profiles for *ever* and *any* may reflect effects of the linear positions of the two NPIs on memory dynamics: subject-verb binding reactivates the correct main clause subject immediately prior to *any*, giving this subject an activation boost that reduces intrusion from the inappropriate licensor; no such advantage is available for pre-verbal *ever*. These results show that the illusory licensing effect observed for the NPI *ever* does not extend to *any*. Further, the sensitivity to the principles of memory dynamics is compatible with a memory retrieval account of NPI licensing in which retrospective dependency-formation is differentially susceptible to intrusion, as a consequence of independent retrieval operations.

(1)   a. Appropriate Licensor: No bills that the senators voted for will ever become law.
       b.InappropriateLicensor: *The bills that no senators voted for will ever become law.
       c. No Licensor: *The bills that the senators voted for will ever become law.
(2)   {**No/The**} criminals that {**the/no**} policemen could catch have felt **any** satisfaction from a crime.
(3)   {**No/The**} criminals that {**the/no**} policemen could catch have **ever** felt satisfaction from a crime.

References. [1] Vasishth et al. 2008. *Cog. Sci.* [2] Xiang et al. 2009. *Brain Lang.* [3] Kadmon & Landman. 1993. *Ling. Philos.* [4] Giannakidou. 2011. Semantics. [5] Lewis & Vasishth. 2005. *Cog. Sci.*

Poster Session 2

Information structure and the 'height' of ellipsis

Timothy Dozat (Stanford University) & Jeffrey Runner (University of Rochester)
tdozat@stanford.edu

Background. It has been widely noted in the literature that VP ellipsis sentences are degraded by voice mismatch (Hankamer and Sag, 1976; Tanenhaus and Carlson, 1990; Frazier and Clifton, 2006; Arregui et al., 2006; Kim et al., 2011), although to varying degrees (as noted by Dalrymple et al. (1991); Hardt (1993), inter alia). Merchant (2008) compares mismatch in VP ellipsis to mismatch in the related pseudogapping construction, noting that the latter seems to be categorically unacceptable, and proposes that the constructions differ in the node which deletes: in VPE the node below the [VOICE] feature is elided, allowing for voice mismatch; but in pseudogapping the node above the [VOICE] feature is elided, so there is a lack of structural parallelism if the [VOICE] features disagree. This 'height of ellipsis' account extends to sluicing (see Tanaka (2011); San Pietro et al. (2012)), which elides an entire TP/S, along with its [VOICE] feature, predicting categorical ungrammaticality in voice mismatched sluicing sentences (compare (1a-1b)). What Merchant's proposal lacks, however, is an account of why some voice mismatched VP ellipsis sentences are unacceptable, as noted above (compare (2b-2d)). Kertz (2010) observes that the degree of mismatch is related to information structural features of the two clauses: "contrastive topic" VP ellipsis has a strong bias to contrast the subject of the elided VP with the subject of the antecedent clause; voice mismatch disrupts this preference for syntactic parallelism (compare (2a-2b)). However, in "simple focus" VP ellipsis, the contrast does not involve the arguments of the two clauses so parallelism is no longer necessary (2c-2d).

Experiment. We compared voice mismatch in both forms of VP ellipsis and in sluicing. Participants (n=29) read sentences like (1-2) and rated their acceptability on a "thermometer" scale (Featherston, 2009). Half the materials included the struckout material as unelided controls. Since sluicing elides everything except the wh-phrase (including the [VOICE] feature), Merchant's analysis, which makes no distinction between the two types of VP ellipsis, would seem to predict that sluicing mismatches should be worse than VP ellipsis mismatches. However, because sluicing focuses the wh-phrase of the ellipsis clause, Kertz's analysis would predict that contrastive topic VP ellipsis sentences should be about as bad as sluicing sentences, but simple focus VP ellipsis should be significantly better.

Results. Our data were mostly consistent with Kertz's analysis: simple focus VP ellipsis mismatches were better than corresponding contrastive topic ($p < 0.001$) sentences and roughly equal to corresponding sluicing ($p > 0.05$) sentences (but numerically higher), while contrastive topic VP ellipsis mismatch sentences were actually worse than sluicing mismatch sentences ($p < 0.005$).

Discussion. Our study contributes to the growing body of research showing the roles of discourse and information structure on the interpretation and processing of anaphoric expressions (Kehler, 2000; Kaiser and Trueswell, 2004; Winkler, 2000; Hendriks, 2004; Kertz, 2010, inter alia), by extending it to sluicing.

1) SLUICING: Someone read *The Lord of the Rings* for class,
    a) MATCH: but I don't know who ~~read it~~.
    b) MISMATCH: but I don't know by whom ~~it was read~~.
2) VPE: Cindy read *The Lord of the Rings for class,*
    a) CT MATCH: but I don't know if Elizabeth did ~~read it~~, too.
    b) CT MISMATCH: but I don't know if *The Once and Future King* was ~~read by her~~, too.
    c) SF MATCH: but I don't know if she had to ~~read it~~.
    d) SF MISMATCH: but I don't know if it had to be ~~read~~.

Here comes the subject: Listeners use number-marked verbs to predict subject number

Cynthia Lukyanenko & Cynthia Fisher (University of Illinois, Urbana-Champaign)
lukyane1@illinois.edu

During sentence comprehension, we integrate lexical, syntactic, and discourse or referential context to anticipate upcoming material. For example, listeners hearing *"The girl will taste…"* look quickly to candy rather than a beer or a carousel, taking into account both world knowledge about the already named agent and the semantic constraints of the verb [1]. Listeners also use morphosyntactic information, such as tense and case, to direct looks to relevant parts of a scene. For instance, hearing *"The man will drink.."* listeners direct looks to a full coffee cup, not an empty beer glass [2].

In two experiments, we investigated listeners' ability to anticipate the grammatical number of an upcoming noun from a number-marked verb in questions and locative inversions (e.g., *Where <u>are</u> the good <u>cookies</u>?, Here <u>comes</u> the pretty <u>bunny</u>!*). The agreement marking on verbs is a purely syntactic cue that does not modify the meaning of the sentence, but serves as a pointer connecting the verb to its subject. Will listeners use it predictively in sentence processing, or wait for the more substantive information carried by the noun? In Experiment 1, participants saw pairs of pictures, accompanied by sentences of two types: Informative sentences included an agreeing verb before the noun (*Where <u>are</u> the good cookies?*), and Uninformative sentences did not (*Look at the good cookies!*). The determiner and an uninformative adjective formed a brief pre-noun interval in which we could measure anticipatory looks to the upcoming target. For an experimental group (n=24), on each trial one picture showed a single object, and the other showed a pair of identical objects (e.g., one apple, two cookies). A control group (n=24) heard the same sentences, but each pair of pictures matched in number (two apples, two cookies); rendering the "informative" sentences uninformative in context. Visual fixations to the pictures as sentences unfolded revealed that listeners used agreeing verbs to anticipate the number of the upcoming subject. In the pre-noun analysis window, listeners were more likely to make saccades from distractors to targets, and looked longer at the target, in informative than in uninformative trials. The control group showed no such advantage. This effect emerged similarly for regular and irregular verbs. Did listeners anticipate the notional number (e.g., 2 items) or the grammatical number (e.g., plural-marked noun) of the upcoming subject? In Experiment 1, notional and grammatical number coincided: plural NPs referred to multiple things and singular NPs to single objects.

In Experiment 2, we asked whether listeners used a number-marked verb to anticipate the grammatical number of the upcoming subject. We did so by varying the category of the subject nouns (count, mass, invariant plural), while holding constant the number of objects in the paired pictures. Thus, participants (n=16) heard sentences paired with pictures as in (1-2), below. Participants again showed an advantage in informative relative to uninformative trials, looking reliably more at the target than the distractor in the pre-noun window.

(1) <u>Mass</u>   Inf: "Where is the good toast?"   [2 bananas, 2 pieces of toast]
                     Inf: "Where are the good bananas?"   [2 bananas, 2 pieces of toast]
                     Uninf: "Where is the good toast/banana?"   [1 banana, 1 piece of toast]

(2) <u>Invar. Pl.</u>   Inf: "Where are the pretty glasses?"   [1 phone, 1 pair of glasses]
                     Inf: "Where is the pretty phone?"   [1 phone, 1 pair of glasses]
                     Uninf: "Where are the pretty glasses/phones?"   [2 phones, 2 pairs of glasses]

Thus, listeners use the information carried by a number-marked verb to anticipate the number features of an upcoming noun, without relying solely on number meaning. Even semantically poor morphosyntactic cues like verb number agreement are rapidly integrated and used to anticipate upcoming material during online sentence comprehension.

[1] Kamide, Y., Altmann, G., & Haywood, S. (2003). *Journal of Memory and Language*
       [2] Altmann, G., & Kamide, Y. (2007). *Journal of Memory and Language*

Poster Session 2

An ACT-R model interfacing eye movements with parsing

Felix Engelmann, Shravan Vasishth, Ralf Engbert, & Reinhold Kliegl (University of Potsdam)
felix.engelmann@uni-potsdam.de

We present a model in the cognitive architecture ACT-R (Anderson et al., Psychological Review, 111[4], 2004) that combines oculomotor control in reading with post-lexical processing by interfacing the eye movement model EMMA (Salvucci, Cognitive Systems Research, 1[4], 2001) with the cue-based retrieval parser developed by Lewis & Vasishth (Cognitive Science, 29[3], 2005).

The way parsing difficulty interacts with eye movement control is inspired by E-Z Reader 10 (Reichle et al., Psychonomic Bulletin & Review, 16[1], 2009). (1) "Slow integration failure" or time out: If structural integration of a word is not finished by the time the recognition of the next word completes, a short one-word regression is programmed. The function of this regression is to buy time for the integration to finish before the integration of the next word. (2) "Rapid integration failure": When integration fails due to a wrong attachment earlier in the sentence (e.g., in a garden path), the integration process is canceled and again, a regression is initiated. This regression, however, does not serve as a time out. Empirically, the detection of an attachment error leads to an active repair attempt, often observed as long-range regressions and re-reading of earlier material. Thus, regressions resulting from rapid integration failure can target earlier parts of the sentence than just the previous word.

The extended EMMA model was evaluated on the Potsdam Sentence Corpus (Kliegl et al., European Journal of Cognitive Psychology, 16[1], 2004). The duration of each word's post-lexical processing was defined using pre-calculated values for memory retrieval and syntactic surprisal (Hale, NAACL, 2001; Levy, Cognition, 106[3] 2008) generated by Boston et al. (Eye Movement Research, 2[1], 2008; Language and Cognitive Processes, 26[3], 2011). The evaluation assessed the model's prediction of a range of fixation measures including regression rates, grouped by classes of word frequency. In addition to six first-pass measures that eye movement models are usually tested on (gaze, first fixation, and single fixation duration, and the rate of skipping, fixating once, and fixating twice), we included total fixation time, regression-path duration, re-reading time, re-reading rate and first-pass regression rate. Retrieval and surprisal were used individually or combined; Surprisal was also tested as directly affecting word recognition. Out of six tested configurations, the best model involved retrieval and surprisal jointly affecting the duration of word integration. The root-mean-square error between data and predictions decreased from 0.638 for the pure EMMA model to 0.206 for the extended model. Moreover, a linear regression showed that the model predicts positive effects of surprisal and retrieval for most fixation measures, confirming an empirical corpus study by Boston et al. (2011).

Our evaluation implicates that a combination of memory retrieval and structural expectation is a good model for post-lexical syntactic processing in sentence comprehension. Based on this conclusion, we propose a psychologically realistic serial parsing variant of surprisal, which can be implemented within the ACT-R parser. This proposal links surprisal with parsing difficulty through delayed firing of ACT-R production rules. Instead of maintaining parallel parses, here surprisal is determined by trained preferences for parsing rules at each parsing step. After every word integration cycle, individual firing latencies of all parsing rules are adjusted as a function of their probability of being selected in the next cycle. These probabilities theoretically represent the reader's expectations learned through experience. Lexical input then constrains the set of applicable rules and, in case of a "surprising" word, might exclude highly expected parsing steps. In this case, the parser has to choose among non-preferred, slower executing rules, which slows down integration. This account grounds surprisal in procedural knowledge, that is learned through reading experience. It builds on the intuitive notion that trained processes are carried out faster.

In combination with the proposed surprisal implementation, the EMMA/ACT-R reading model would be the first model explicitly formulating the interaction of memory retrieval and syntactic expectation with eye movement control.

Form-based syntactic expectations affect the duration of early fixations in reading
Thomas A. Farmer (University of Iowa), Klinton Bicknell (University of California—San Diego) & Michael K. Tanenhaus (University of Rochester)
thomas-farmer@uiowa.edu

A growing literature suggests that readers and listeners actively generate predictions about various aspects of incoming linguistic input. But do expectations from context map onto lower-level expectations, and if so, how? We propose that comprehenders use internally generated predictions at multiple levels to explain the source of the input (A. Clark, *in press*, *BBS*). A stream of hierarchically organized generative models propagates higher-level expectations to lower-level (progressively closer to perceptual cortex) models via feed-backward connections. At each level, as the incoming signal is intercepted, form-to-expectation mismatch produces prediction error, which is then fed forward into the system so that higher-level models can be updated to better approximate the input at time t+1.

Context should, therefore, influence perceptual processes occurring very early in the processing stream. Farmer, Christiansen, and Monaghan (2006, *PNAS*) demonstrated that when a sentential context conferred a strong expectation for a word of a given grammatical category (as in *The child saved the …* , where a noun is strongly expected), subjects were slower to read the incoming noun when its form (i.e., its phonological/orthographic properties, quantified by a feature-based measure of word-form typicality detailed in Farmer et al., 2006) was atypical with respect to other words in the expected category. In a subsequent MEG experiment, Dikker, Rabagliati, Farmer, & Pylkkanen (2010, *Psych. Sci.*) showed that at about 100ms post-stimulus onset—timing that is unambiguously associated with perceptual processing—a neural response was elicited when there was a mismatch between form and syntactic expectation. Moreover, the source of the effect was localized to the occipital lobe, consistent with the hypothesis that form-based syntactic expectations are influencing early perceptual processing. If the form-based hypothesis is correct, then effects should be observed in eye-movement measures that have been argued to reflect primarily early lexical processing, a prediction that was not supported by results of Staub, Grant, Clifton, and Rayner (2009, *JEP:LMC*). However, the design of Staub et al.'s studies encouraged adaptation/learning effects that would have reduced or eliminated form-based expectations (Farmer, Monaghan, Misyak, & Christiansen, 2011, *JEP:LMC*).

We report an eye-movement experiment that tests the form-based hypothesis. Plausibility, frequency, and length were all controlled between conditions. In Experiment 1, subjects read sentences in which the word-form properties of the target word were either congruent ('noun-like' noun, e.g., marble) or incongruent ('verb-like' noun, e.g., insect) with a context conferring a strong noun expectation (ex. 1a-b). Mixed-effects models revealed strong expectation effects: first-fixation duration, gaze duration, go-past time, and total reading time were all longer for verb-like nouns in strong noun expectation contexts. In an ongoing experiment, expectation strength was manipulated by placing half of the words in a less N-biased context (ex. 2), with the remaining half of the words left in the biasing context. Preliminary analyses reveal a marginal interaction between context and word-form typicality on gaze duration: a typicality effect was observed in the strong N-bias context but not in the less-biased context.

These experiments provide novel eye-movement evidence that linguistic context is used to generate perceptual expectations about form-based properties of upcoming words during reading. Moreover, form-to-expectation matching, and model-based updating based on error signals early in the processing stream, may play a central role in readers and listeners' ability to rapidly evaluate a linguistic signal under time pressure.

Example Sentences

(1a) The curious young boy saved the marble that he … (Noun-like Noun, strong N-biased context)
(1b) The curious young boy saved the insect that he … (Verb-like Noun, strong N-biased context)
(2) The teacher wrote the word "marble" / "insect" on the blackboard. (Less N-biased context)

Poster Session 2

Electrophysiological response to manipulation of syntactic expectations
Joe Kirkham, Chelsea Guerra & Edith Kaan (University of Florida)
FDLKirkham@gmail.com

There is ample evidence that readers anticipate upcoming information when processing sentences (e.g., Kutas et al., 2011). Exactly how and when syntactic predictions affect the processing of incoming words is still under debate. According to Lau, Stroud, Plesch, and Phillips (2006), syntactic predictions can modulate the ELAN effect for ungrammatical continuations of a sentence: a larger ELAN effect was found when the critical word violated a strong syntactic prediction, relative to a condition in which no clear syntactic prediction could be made. Results from MEG studies (e.g., Dikker et al., 2009), suggest that syntactic predictions can modulate the sensory responses to the word form of the ungrammatical continuation. The aim of the current ERP study was to test the effect of structure-based predictions as well as experimentally-induced predictions on the processing of word category violations.

The current study built upon the experiment conducted by Lau et al. (2006). EEG was recorded from 20 native English speakers while they completed a grammaticality judgment task in which they read sentences that varied in degree of expectancy of the critical position (see (1), critical position underlined).
1. a/b. Although John met Max's surgeon, he did not meet Bill's {*of / before} the operation. (ellipsis possible)
   c/d. Although the surgeon met Max, he did not meet Bill's {*of / *before} the operation. (ellipsis impossible)
In (a) and (b), the sentence can be completed at *Bill's*, whereas in (c) and (d) an overt noun is expected to follow the possessive. In contrast to Lau et al. (2006), the paradigm included sentences of the type (b, grammatical) and (d, ungrammatical). In addition, the critical word in the (a) and (c) condition was always *of;* whereas in (b) and (d) *before*, *during* and *after* each occurred one-third of the time, allowing us to investigate the effects of experimentally-induced expectations. ERP analysis was conducted on trials to which the participants responded correctly.

Contrary to Lau et al. (2006), there was no ELAN effect at the critical word in the (c) vs.(a) comparison, even when the ERP signal was re-referenced to the average of all electrodes (as in Lau et al., 2006). This is not surprising, given the inconsistency of the ELAN in the literature (e.g., Steinhauer and Drury, 2012). Instead, a significantly larger, right-lateralized fronto-central negativity was found around 300-500ms for the critical word in (a) compared to (c). In addition, the critical word in (c) showed an earlier and larger P600 component than in (a). Since the combination of possessive+*of* was always ungrammatical, the differences seen between the (a) and (c) conditions cannot be due to differences in experimentally-induced expectations, but are likely due to differences in linguistic expectations based on the preceding structure: in (c) an overt noun is expected, but this expectation is not met, leading to an earlier positivity, whereas in (a) an elided noun can be licensed, leading to a 300-500ms negativity. However, no significant differences were found in the ERPs between the (b) and (d) conditions, in which the specific critical word varied and therefore was not highly expected. This suggests that the processing of the ungrammatical continuation is affected by linguistically-based predictions only if the specific critical word is highly expected. In contrast to Dikker et al. (2009)'s findings we did not find a modulation of early sensory responses for (a,c) versus (b,d), suggesting that the interaction of linguistically- and experimentally-induced expectation occurs at a higher level of processing.

References
Dikker, S., Rabagliati, H., Pylkkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition, 110*, 293-321.
Kutas, M., DeLong, K. A., Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In M. Bar (Ed.), *Predictions in the brain: Using our past to generate a future* (pp. 190–207). OUP
Lau, E., Stroud, C., Plesch, S., Phillips, C. (2006). The role of structural prediction in rapid syntactic analysis. *Brain and Language*, *98*, 74-88.
Steinhauer, K., and Drury, J.E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language*, *120*, 135-162.

Effects of verb bias and syntactic ambiguity on reading in people with aphasia
Gayle DeDe (University of Arizona)
gdede@arizona.edu

People with aphasia often have sentence comprehension impairments. According to the Lexical Bias Hypothesis (*LBH*), sentence comprehension impairments in people with aphasia emerge when the verb's argument structure bias conflicts with the sentence structure (Gahl, 2002). Recent work suggests that this hypothesis can account for comprehension deficits that affect structurally simple sentences, such as active transitive and intransitive sentences. However, there has not been a clear demonstration of how verb bias affects comprehension of syntactically ambiguous sentences for people with aphasia. The purpose of this study was to test how verb bias affects comprehension of syntactically ambiguous and unambiguous sentences using an on-line measure of written sentence comprehension, self-paced reading.

Nine people with aphasia (4 Anomic, 2 Broca's, 2 mixed non-fluent, 1 Conduction) and 10 age-matched controls performed a self-paced reading task on a subset of the sentences from Garnsey, Pearlmutter, Myers, and Lotocky (1997). Seven of the people with aphasia exhibited impaired sentence comprehension on the Northwestern Assessment of Verbs and Sentences. All of the critical sentences contained sentential complements (ex. 1 & 2). The main verb was biased to take a sentential complement (SC) or a direct object (DO). The sentential complement was either marked by the complementizer *that (*i.e., an unambiguous sentence) or unmarked (i.e., an ambiguous sentence). The ambiguous noun phrase (*the fire/ the airplane*) was always an implausible object for the main verb. All items were followed by a yes/no comprehension question.

1. The talented photographer| accepted| (that)| <u>the fire</u>| <u>could not</u>| have been| prevented. (DO)
2. The ticket agent| admitted| (that)| <u>the airplane| had been</u>| late| taking off. (SC)

Because all of the critical sentences contained sentential complements, the *LBH* predicts that people with aphasia will experience greater processing disruptions when the verb is DO-biased. The effects of syntactic ambiguity were more difficult to predict because people with aphasia are known to have difficulty processing function words.

Residual reading times (RTs) were analyzed to control for differences in segment length. The critical segments (ambiguous noun phrase and disambiguating region) are underlined in (1) and (2). **Control Group:** RTs were significantly longer in *no-that* vs. *that* sentences at both the ambiguous and disambiguating segments. The ambiguity effect was numerically larger in sentences with DO-biased verbs, but the interaction was not significant at either segment. **People with aphasia:** At the ambiguous segment, the interaction between verb bias and ambiguity was significant in the analysis by participants but not items. Inspection of the means suggested that RTs were longer for DO-, but not SC-, biased verbs in the *no-that* vs. *that* condition. RTs at the disambiguating segment were significantly slower for sentences with DO- than SC- biased verbs. Although the effect of verb bias was numerically larger in *no-that* sentences, the same pattern was present in *that*-sentences and the interaction was not significant. Inspection of the individual case data showed that 7 of the 9 participants with aphasia showed a larger effect of verb bias than ambiguity at the disambiguating region.

In general, the results were consistent with the *LBH.* The control group's RTs at both the ambiguous noun phrase and the disambiguating region seemed to be most influenced by whether or not the sentence was ambiguous. Taken together, the people with aphasia's RTs for the two critical segments suggested that they were sensitive to the structural cue provided by *that*, but relied on verb bias to a greater extent. The over-reliance on verb bias in people with aphasia may be due to delayed or otherwise impaired processing of closed-class words.

The role of the left anterior temporal lobe in semantic memory vs. sentence processing

Masha Westerlund[1], Doug Bemis[2], & Liina Pylkkänen[1] ([1]New York University, [2]CEA-INSERM Neurospin)
masha.westerlund@nyu.edu

Introduction: A growing body of cognitive neuroscience literature focusing on combinatorial operations in language has robustly implicated the left anterior temporal lobe (LATL) in the processing of sentences [1] and of basic phrases [2]. At the same time, the LATL has also been implicated in semantic processing by a separate body of work focusing on semantic dementia and the retrieval of specific compared to basic concepts [3,4]. The similar localization and intuitive overlap between these results raise an intriguing possibility: there may in fact be a shared mechanism whereby the LATL participates in the specification of concepts within and across words. For example, the LATL may function as a 'semantic hub' that binds features of concepts across a distributed network [4]. The current experiment aimed to investigate this possibility. We independently manipulated composition and concept specificity within a minimal MEG paradigm in which participants viewed simple nouns that denoted either less specific level (*fish*) or more specific categories (*trout*) presented in either combinatorial (*blue fish/trout*) or non-combinatorial contexts (*xhsl fish/trout*). If the prior composition and specificity findings reflect a shared mechanism, we would expect similar main effects of both specificity and composition in the LATL.

	Low Spec		High Spec	
Combinatory (two words)	blue boat		blue canoe	
Non-combinatory (one word)	qktz boat		qktz canoe	

Design: We used a task from the literature on specificity [4] in which subjects had to match images to either less specific ('boat) or more specific nouns ('canoe'), during MEG recordings. Crucially, the nouns were either preceded by a consonant string (e.g. 'xhsl boat'/'canoe'), or in a combinatory context (preceded by an adjective, e.g. 'blue boat'/'canoe'). Though it could be argued that a comparison to a consonant string condition might be confounded by the presence of extra lexical material in the two word condition, previous results have shown that a composition effect is not elicited when two non-combining words are compared to a word preceded by a consonant string [2], suggesting that this comparison does in fact capture composition operations rather than simply increased lexical processing. Trials were presented word-by-word, and neural activity was measured from the onset of the target noun in all conditions, such that the activity elicited by combinatory and non-combinatory operations was measured at the same lexical items. The target nouns were matched for log frequency, length, and lexical decision reaction time from the English Lexicon Project, and the adjective-noun combinations for transition probability.

Results: Instead of straightforward main effects of composition and specificity, a non-parametric cluster-based analysis [5] identified an interaction between these two variables: significantly greater composition-related activity in the LATL for the less specific nouns ('blue boat' v. 'xhls boat') as compared to the more specific nouns ('blue canoe' v. 'xhls canoe') between 222-254 ms, corresponding to the timing of previous MEG effects of composition in the LATL [2].

Conclusion: Though these results do not point to a straightforward account of a shared mechanism, they show an intriguing interaction between composition and specificity-related activity. This result suggests that the two mechanisms do not operate entirely independently within the LATL, and therefore demonstrates the need for further enquiry into the relationship between semantic memory and composition.

[1] Stowe et al. (1998). Neuroreport, 9, 2995-2999. [2] Bemis & Pylkkänen (2011), J. Neuroscience, 31, 2801-2814. [3] Tyler et al. (2004). JCN, 16, 351-363 [4] Rogers et al. (2006). J. of Cognitive, Affective, & Behavioral Neuroscience, 6, 201-213. [5] Maris & Oostenveld (2007). J. of Neuroscience Methods, 164, 177-190.

Semantic similarity-based competition in sentence production and comprehension
Gina Humphreys (University of Manchester) & Silvia Gennari (University of York)
silvia.gennari@york.ac.uk

Neurocognitive models of language suggest that production and comprehension share common syntactic-semantic sentence-level processes in pars opercularis (PO) within left inferior frontal gyrus (LIFG) (Menenti et al., 2011). However, psycholinguistic models suggest different component processes across tasks, rather than commonalities (e.g. retrieval vs. recognition). Here we investigate how similarity-based competition affects behavioral and brain responses in order to examine shared or distinctive sub-processes in each task and their implications for these models.

It has been shown that relative clauses (RCs) with two animate nouns are described with passives like (1a), whereas RCs with nouns of different animacy are described with active RCs [2]. It's argued that during planning, animate entities compete with each other due to high similarity resulting in the demotion of one entity to sentence-final position. Here, in both production and comprehension, we manipulated semantic similarity – high vs. low similarity as in (1a,b) and (2a,b)– and used the picture-based RC elicitation task in Gennari et al., 2012, and a RC reading task after seeing a related picture. Thus, both tasks were grounded on the same visual image.

Results indicated that the degree of noun-noun similarity correlated with the proportion of passive RCs produced and the reading times of actives (1b)-(2b), which have the same syntax. Thus, semantic competition affected both tasks but in different measures (structure choices vs. reading times).

(1a) The boy being kicked by the girl. (1b) The boy that the girl is kicking. (high-similarity)

(2a) The boy being bitten by the dog. (2b) The boy that the dog is biting. (low-similarity)

Using the same tasks and manipulation, an fMRI study investigated the brain regions recruited in production and comprehension. The results showed that semantic competition modulated neural responses within PO in both tasks, however production alone engaged an area associated with motor planning and decisions, the supplementary motor area (SMA) (Figure 1). Interestingly, despite some common PO activity, production engaged more dorsal PO areas, whereas comprehension engaged more ventral aspects of PO (Figure 2).

Together, the results suggest that semantic competition affected lexical selection and positional (sequencing) processes in production, as revealed by word order choices (passives) and the recruitment of motor-related SMA. In comprehension, it affected thematic role assignment, as revealed by RTs and the involvement of PO, previously associated with such processes. Interestingly, these processes appear segregated within PO, suggesting distinctive competition resolution mechanisms for each task. Overall, production and comprehension remain distinct as suggested by psycholinguistic models, despite some common processes, which we will argue, reflect the management of shared linguistic knowledge.

Figure 1
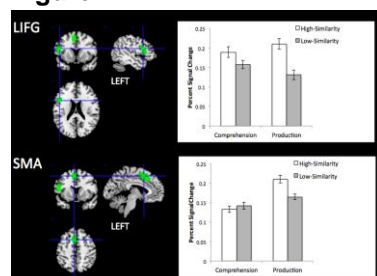
Figure 2
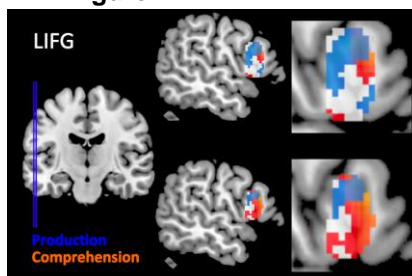
References

Menenti, L., Giehan, S., Segaert, K., Haagort, P. (2011*). Psych. Sci., 22*, 1173-82.
Gennari, S., Mirković, J., MacDonald, M. (2012*). Cognitive Psychology, 65*,41-176.

Rethinking the functional significance of early negativity

Lisa Rosenfelt, Robert Kluender & Marta Kutas (UC, San Diego)
lrosenfelt@ucsd.edu

Neurocognitive processing models [1] often assume a modular architecture [3] and thus map linguistic ERP component latency onto discrete stages of serial parsing. The linchpin of these models is the early negativity (EN), elicited between 100-300 ms by word category violations (WCVs) [4,5] that interrupt structure-building operations, and thus often mapped onto an automatic and informationally encapsulated first stage of syntactic parsing [6].

However, recent studies have shown the link between EN and WCVs to be tenuous: EN is elicited by a WCV only when (a) it occurs in a highly predictive syntactic context [7], (b) the probability of violation is 50% or lower [8], (c) preceding context is held constant [9] or (d) sensory (visual or auditory) word form features differ greatly from those of the expected category [10-12]. In view of such results, an alternative hypothesis is that EN is elicited by violations of expectation of sensory form in highly predictive contexts [11-14].

In three studies involving five stimulus sets {1-5} (see Table1), we have now shown that a WCV is neither sufficient {2,3,5} nor necessary {4} for eliciting EN during sentence processing. Taken together, our results undermine assumptions of serial neurocognitive processing models in the following ways: (a) EN is elicited by WCVs in only one syntactic context in our studies {1} but has twice been elicited by plausible semantic substitutions in idioms {4}; (b) WCVs consistently elicit an N400/P600 complex {1-5}. In other words, we have found that contextually driven semantic/pragmatic processing elicits an early response while the processing of syntactic structure-building consistently elicits late responses.

However, sensory form-based expectation accounts were also partially falsified by our results: visually unexpected plausible nouns elicited EN {4} while the equally unexpected closed-class WCVs did not {5}. Further studies are clearly needed to narrow down the functional significance of the EN, as the lack of an EN effect in {5} runs counter to predictions of both word category information and sensory form-based expectation accounts.

Table 1.

Violation	EN	N400	P600
1. WCV (unexpected onset): *men's OF shouts*	X	X	X
2. WCV (+ unexpected suffix): *he immediately ALTERATION*		X	X
3. WCV (- expected suffix): *make a closer INSPECT*		X	X
4. Semantic (unexpected onset): *Don't rock the CANOE*	X	X	X
5. WCV (unexpected onset)): *Don't rock the HENCE*		X	X

References. [1] Friederici (2002,2004) [3] Fodor (1983)) [4] Neville et al. (1991) [5] Friederici, Pfeifer & Hahne (1993) [6] Inoue & Fodor (1995) [7] Lau et al. (2006) [8] Magnuson et al. (CUNY 2010) [9] Steinhauer & Drury (2012) [10] Rosenfelt et al. (CUNY 2009) [11] Dikker et al. (2009) [12] Dikker et al. (2010) [13] Dikker & Pylkkänen (2011) [14] Rosenfelt et al. (AMLaP 2011)

Patient studies of language in the modern era

Julius Fridriksson (University of South Carolina)
fridriks@mailbox.sc.edu

Much of the early knowledge about language organization in the brain came from lesion studies of patients with localized brain damage and different kinds of language impairment. Work by Broca (1865), Wernicke (1874), and Lichtheim (1885) towards the latter part of the 19th century not only set the tone for modern studies of brain-language relationships but also continues to have more broad influence on current research in neuropsychology. However, the advent of modern neuroimaging and neurophysiological techniques has placed less emphasis on patient studies for understanding how the normal human brain supports language and speech processing. One of the most influential models of speech and language in modern neurolinguistics is the dual stream model (Hickok & Poeppel, 2007; Poeppel et al., 2012). Although some of the data that motivate this model were derived from studies in normal subjects, it is clearly influenced by the early work on language impairment in patients. I will argue that current patient studies can shape contemporary neuropsychological models of speech and language, especially when combined with modern techniques used to study brain function. Along with other patient studies, we have shown that the temporal lobe plays a crucial role in grammatical processing (Magnusdottir et al., 2012). In a series of experiments in patients with neurodegenerative disease, Wilson and colleagues (e.g. Wilson et al., 2011; Wilson et al., 2012; Deleon et al., 2012) suggest that specific white matter connections between anterior (e.g. Broca's area) and posterior (temporal lobe) language regions via the dorsal stream are crucial for syntactic processing. By contrast, some of our most recent work suggests that the dorsal stream probably has no greater importance for syntactic processing than more ventral connections (via the uncinate fasciculus) between anterior and posterior language regions. Finally, if time allows, I will discuss some of my most recent work showing that patients with non-fluent speech and, in some cases, agrammatism, can produce fluent (and grammatically correct) speech with the aid of external audio/visual feedback and what this could mean for normal speech and language processing.

Sentential context modulates early phases of visual word recognition: Evidence from a training manipulation

Vicky T. Lai (Max Planck Institute for Psycholinguistics), Albert Kim (University of Colorado, Boulder), & James M. McQueen (Radboud University Nijmegen; Max Planck Institute for Psycholinguistics)
Vicky.Lai@mpi.nl

How does sentential context influence visual word recognition? Recent neural models suggest that single words are recognized via a hierarchy of local combination detectors [1]. Low-level features are extracted first by neurons in V1 in the visual cortex, features are then combined and fed into the higher level of letter fragments in V2, and then letter shapes in V4, and so on. A recent EEG study examining word recognition in context has shown that contextually-driven anticipation can influence this hierarchy of visual word recognition early on [2]. Specifically, a minor mismatch between the predicted visual word form and the actual input (*cake* vs. *ceke*) can elicit brain responses ~130 ms after word onset [2].

The current study further investigates how and when conceptual-semantic knowledge influences visual word recognition. 36 participants took part in a training session followed by a testing session; EEG was recorded throughout. Instead of using nonwords like "*ceke*" as in [2], we used an ambiguous symbol ∂, pretested to be midway between 'a' and 'o'. During training using a perceptual learning paradigm [3], half of the participants were trained to see ∂ as a (they saw ∂ in lexically-biased a-contexts like *l∂ke*). The other participants learned that ∂ was 'o' (seeing it in lexically-biased o-contexts like *d∂g*). During testing, all participants saw a-o minimal-pair words (c∂ke for *cake/coke*) embedded in highly predicting sentences (cloze = 80%). Depending on training, contextually-driven anticipation could be congruent or incongruent with the input. That is, "*bake a c∂ke*" would be incongruent for the o-group because "*bake a coke*" is senseless, but congruent for the a-group. If semantic anticipation can influence an early phase of processing, we expect brain responses at ~130 ms [2].

Training was successful. Participants accurately identified ∂-words as words (91%). During training, ∂-words elicited a larger N170 component (160-180 ms) than unambiguous words at a right occipital-temporal site (Fig 1), followed by N400 and P600 effects at the central-parietal sites. During testing, ∂-words in the incongruent condition elicited a larger occipital-temporal P1 (115-135 ms) component than the same ∂-words in the congruent condition (Fig 2), with no N170 or later effects.

These data provide evidence for rapid use of conceptual-semantic knowledge during the recognition of a newly learned letter in context. When the o-group read "*bake a c∂ke*", for example, the contextually supported representation (CAKE) allows anticipation of the "a", which conflicts with the training-supported interpretation (COKE) of the input (c∂ke). Because the ambiguous letter is identical across conditions, the conflict appears to be with the interpretation of the ambiguous letter and not with its physical features, hence specifying the timing of the letter interpretation stage [1] in the processing hierarchy.

Figure 1. Training phase, single words
- real a or o words
- words with a/o symbol

Figure 2. Testing phase, sentence embedded minimal pairs
- congruent
- incongruent
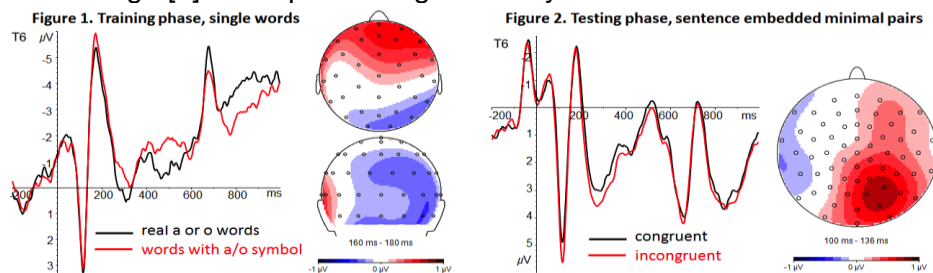
[1] Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. Trends in Cognitive Sciences, 9, 335–341.
[2] Kim, A., & Lai, V. T. (2012). Rapid interactions between lexical semantic and word form analysis during word recognition in context: Evidence from ERPs. Journal of Cognitive Neuroscience, 24, 1104-1112.
[3] Norris, D., Butterfield, S., McQueen, J. M., & Cutler, A. (2006). Lexically guided retuning of letter perception. Quarterly Journal of Experimental Psychology, 59(9), 1505-1515.

Partially activated words participate in combinatory semantic interpretation during sentence processing

Sarah Johnstone, John Trueswell & Delphine Dahan (University of Pennsylvania)
sarahj@psych.upenn.edu

Understanding the interactions between gradient, continuous properties of speech and the more symbolic aspects of linguistic structure can inform our understanding of language processing and of cognition more generally [1]. The present work examines this issue by considering the interface between probabilistic lexical processing and compositional, sentence-level meaning. We demonstrate that partially activated words are subject to combinatory semantic interpretation during sentence processing. Specifically, we show that when listeners hear a sentence like "He will shake the leash," they temporarily consider alternative interpretations like "He will rake the leaves," in which the rhyme of the verb and the onset of the direct object noun overlap phonologically with the actual utterance.

In a visual-world eye-tracking experiment, participants heard sentences (e.g., "He will shake the leash") which, on critical trials, were constructed so that the main verb (*shake*) rhymed with another verb (*rake*) and the rhyming verb was predictive of an object noun (*leaves*) starting with the same sounds as the actual object noun (*leash*). Participants viewed pictures of the actual action theme (a leash), its competitor (a pile of leaves), and three distractors. Their task was to click on the object mentioned. In a counterbalanced design, other participants saw the same displays and heard control sentences (e.g., "He will rattle the leash") that matched the critical sentences except that the verb was semantically similar to the critical verb but did not overlap phonologically with a plausible action for the target noun's competitor. If lexical probabilities and semantic processes interact, *leaves* in this example should be a stronger candidate interpretation for "leash" when participants heard "shake" than when they heard "rattle" because they could be constructing an alternate interpretation of the sentence based on some non-zero consideration that the preceding verb had been "rake."

Forty participants completed the visual-world task and 30 more provided typicality ratings for the possible verb-object pairings. These typicality ratings allowed us to adjust for differences in plausibility between, e.g., *shake the leaves*, *shake the leash*, *rattle the leaves*, and *rattle the leash*, that could explain variation in participants' consideration of the target noun and its competitor. We submitted the fixation data to the competitor (e.g., the leaves) to a mixed-effects model using growth curves. We included in the model the each of the verb-noun typicality ratings and a categorical predictor for whether or not the verb (*shake* vs. *rattle*) rhymed with a plausible alternate interpretation (*rake*). Our model showed that competitor fixations over time were in part explained by the typicality of the target noun and competitor given the verb that had been heard. The model additionally revealed an effect of the rhyming status of the verb: Participants directed their attention toward the competitor (the pile of leaves) earlier and more strongly when they heard "shake the leash" than "rattle the leash," suggesting that they were temporarily considering the phrasal alternative "rake the leaves" as a viable candidate for the input even though it received only partial phonological support.

Thus, the way in which phonetic information constrains lexical processing can modulate and be modulated by subsequently derived semantic context. This finding, along with recent work [2], questions the view that phrase-level meaning is built word-by-word only over the outputs of perceptual commitments made at the word level. Indeed, the "right-context" effect demonstrated here suggests instead that probabilistic lexical activation and the combinatorial processes critical to semantic integration are tightly intertwined and can operate over a multi-word window. Future work remains to explore the critical questions of when and how discrete interpretations of the input ultimately emerge out of these ongoing interactions between gradient lexical candidates and phrase-level meaning.

References

[1] Smolensky, Goldrick, & Mathis. (in press). Cognitive Science.
[2] Levy, Bicknell, Slattery, & Rayner. 2009. Proc. Natl. Acad. Sci., 106(50), 21086-21090.

Syntactic priming in comprehension: Priming 'early' closure

Matthew J. Traxler, Megan A. Boudewyn, Tamara Y. Swaab, & Liv J. Hoversten (University of California, Davis
mjtraxler@ucdavis.edu

Two experiments, one involving self-paced reading and one involving event-related potentials (ERPs), investigated readers' responses to sentences containing an "early" vs. "late closure" ambiguity (e.g., 1 and 2):
  (1) While Fernanda was dressing the baby played on the floor.
  (2) John knew Fernanda was dressing the baby.
Sentences like (1) are temporarily ambiguous, because the noun phrase "the baby" could be incorrectly interpreted as the direct object of the verb "dressing" (Frazier & Rayner, 1982; Ferreira, 2003). Sentence (2) is similarly ambiguous because "Fernanda" can be assigned the object role for "knew" (Trueswell et al., 1993). Syntactic disambiguation around the verbs "played" in (1) and "was dressing" in (2) is generally accompanied by increased processing difficulty when sentences like (1) and (2) are contrasted with unambiguous controls, with the magnitude of effects depending on properties of the initial verb.

The current experiments tested whether processing of sentences like (1) and (2) would be facilitated if they were preceded by a syntactically related prime sentence. A self-paced reading experiment involving 48 participants and 48 sets of items tested readers' responses under four conditions. In the first condition, sentences like (1) were preceded by syntactically identical sentences like (3):
  (3) While the cook was dressing the salad fell onto the floor.
In the second condition, an unambiguous "late closure" sentence served as the prime:
  (4) While the cook was dressing the salad the onions fell onto the floor.
The third and fourth conditions had sentences like (4) as targets; with sentences like (1) and (4) as primes. Thus, the experiment crossed two types of primes ("early" closure and "late" closure) and two types of target ("early" vs. "late" closure). Items were rotated across lists so that every prime sentence (from one list) served as a target (on another list). When reading times for the syntactically disambiguating verb (i.e. "fell") and the following word were analyzed, they showed an interaction of prime and target types. Processing of "early" closure sentences was facilitated by the presentation of "early" closure primes, but not by "late" closure primes. Processing of "late" closure sentences was not facilitated by either type of prime. These results indicate that structural processing of "early" closure sentences is facilitated by the immediately prior experience of processing a structurally identical sentence. Semantic overlap, by itself, did not facilitate processing of the sentences at and following the point of syntactic disambiguation.

The ERP experiment tested 24 readers' responses to 80 target sentences like (2). The prime sentences were either structurally identical, like (5)
  (5) The young man knew the answer to the question was in the back of the book.
or structurally mis-matched, like (6)
  (6) The young man knew the answer to the question.
The experiment also included a number of "catch" trials, where sentences like (5) and (6) were followed by structurally unrelated sentences, so that participants could not predict the type of target that would follow. The ERP results showed reduced negativities for target sentences at the first verb (e.g., "knew"). They also showed reduced positivities following the syntactically disambiguating main verb (e.g., "was") in (2). These results, like the self-paced reading results, suggest that processing the syntactic form of the prime sentence facilitates computation of the "early" closure reading of the target.

These results extend the range of sentences for which we have evidence of syntactic priming during on-line comprehension. In addition to these two completed experiments, further experiments are underway to assess the extent to which these outcomes depend on repeating the verb across the prime and target sentences (see also Arai et al., 2007; Carminati & van Gompel, 2009; Thothathiri & Snedeker, 2008a, 2008; Tooley et al., 2009 Traxler, 2008).

Local coherence and digging-in effects in German

Dario Paape, Titus von der Malsburg, Shravan Vasishth (University of Potsdam)
paape@uni-potsdam.de

Tabor, Galantucci & Richardson (2004) showed that subjects reading sentences like "The coach smiled at the player tossed a frisbee" were distracted by the presence of a locally coherent substring ("the player tossed a frisbee") even though English grammar rules this analysis out. The authors claim that this 'local coherence' effect can only be explained by parsing models that compute structures which are incompatible with the previous input. Specifically, 'self-organizing' parsers allow multiple globally incompatible syntactic links to be active simultaneously. Attachment conflicts are resolved through competition between these links. One 'self-organizing' model, SOPARSE (Tabor & Hutchins, 2004), additionally predicts a so-called 'digging-in' effect: attachments grow stronger over time, making it harder for alternative parses to compete. To test this effect, we conducted a self-paced reading study with 40 native speakers of German. Word-by-word segmentation was used. In a 2x2 design, both the availability of a locally coherent structure and the time available to dig in (long vs. short) were manipulated:

Heute weiß man, dass ...
today know one that …

a./b. *einer/einige der Generäle belagerte Städte mit großer Verbissenheit verteidigte(n),*
one/some of the generals besieged cities with great determination defended-(PL)
um Zeit zu gewinnen.
to time to gain.

c./d. **entlang der Nachschubrouten anfänglich** *einer/einige der Generäle*
along the supply lines initially one/some of the generals
belagerte Städte mit großer Verbissenheit verteidigte(n), um Zeit zu gewinnen.
besieged cities with great determination defended-(PL) to time to gain.

'Today we know that (along the supply lines, initially) one/some of the generals stubbornly defended besieged cities to gain time.'

Local coherence was controlled via the number of the subject of the subordinate clause ("one/some of the generals"). In the singular conditions (a, c), the word "besieged" can be interpreted as an adjective in the globally correct SOV structure or as a verb in a locally coherent SVO structure ("the general besieged cities"). SVO word order is ungrammatical in German subordinate clauses headed by "that". In the plural conditions (b, d), no local coherence effect was expected as "besieged" (singular) does not agree with the subject (plural). Digging-in time was controlled by inserting adverbial material at the beginning of the subordinate clause (c, d), which should allow the expected verb-final structure to become entrenched before the locally coherent structure becomes available, thereby reducing competition. On the other hand, a lexical preference to interpret the ambiguous word ("besieged") as a verb rather than an adjective should increase competition by strengthening local coherence. The bias of the critical word towards either reading was calculated from four German newspaper corpora. We analyzed the reading times at each region from the critical word ("besieged") onwards using linear mixed models. The fixed effects were local coherence, digging-in time, lexical bias, and the two-way interactions of these factors. We also added crossed random intercepts and random slopes for all main effects where applicable. In the locally coherent conditions, reading times were faster two words downstream from "besieged" ("with") when the adverbial material was present (c) compared to the case where it was absent (a), confirming the prediction that digging-in reduces competition. When the time for digging in was short (a, b), verb-biased ambiguous words led to increased reading times on "with", suggesting an impact of lexical bias on local coherence. The results indicate that that digging-in and local coherence effects exist in German, and that they interact as predicted by SOPARSE.

Task effects on prosodic prominence

Andrés Buxó-Lugo, Joe Toscano, Duane Watson (University of Illinois at Urbana-Champaign)
buxo2@illinois.edu

Most prosody research is conducted on spoken corpora or on speech samples collected from laboratory studies. However, laboratory studies often use referential communication tasks in which speakers talk to a computer and the communicative "stakes" are low. Previous studies have reported discourse focus effects on duration, intensity, and/or F0 (see Wagner & Watson, 2010, for a review), but these effects are typically small and the distributions of acoustic cues between different discourse contexts (e.g., given vs. new information) are highly overlapping. Thus, these cues may not provide listeners with reliable information about prosodic structure.

This could mean that the specific cues measured are not the best indicators of prosodic prominence or that talkers simply do not reliably convey this information. Another possibility is that prosodic choices are influenced by the degree to which a speaker is engaged in the task. If the task engages their attention, they might provide more informative cues than they would otherwise. If true, prosody use between interlocutors who care about conversational outcomes might look very different from prosody use in laboratory communication tasks. While corpora presumably contain speech from more natural conversations that may exhibit this property, they do not offer the level of control over prosodic context that laboratory tasks do.

We examined this issue by developing two different tasks: a low-engagement task and a high-engagement task. Critically, the words produced on experimental trials in each task are identical. The low-engagement task is similar to the referential communication tasks that are typically used in the field. The high-engagement task is a novel, goal-oriented task set in a computer game. In the game, participants move through a 3D environment and interact with objects in it to solve a series of puzzles. In both tasks, the participant is presented with color sequences that are read aloud. Color sequences are presented in pairs of three colors per trial. In the low-engagement task, the participant sees each color sequence, reads it, and proceeds to the next trial. In the high-engagement (computer game) task, the participant has to communicate with another person over a headset in order to solve a puzzle on each trial and reach the end of the game. Filler trials alternate with critical trials that involve the same color sequences used in the low-engagement task. One participant is presented with the color sequence and conveys that information to the other participant. The second participant enters the color sequence as a "code" in the game, unlocking the door for that trial. Because the same sequences are produced in each task, speakers' productions are comparable across them.

Critically, we manipulated the information status of specific words in each sequence. For both tasks, there were scenarios in which the target words – the second word in the second sequence – were given (already mentioned), new, or carried contrastive focus. Example sentences are in (1) with target words in **bold**:

  (1a): New sequences: red blue green | gray **pink** black
  (1b): Given sequences: gray pink black | gray **pink** black
  (1c): Contrast sequences: gray blue black | gray **pink** black

We analyzed target word duration, F0, and intensity using multilevel models. For duration, we found a main effect of information status, with longer durations for the new and contrast conditions. We also found a main effect of engagement, with longer durations overall for the high-engagement task. Critically, we also found an interaction between task and information status, suggesting that the differences in duration were larger across information status conditions in the high-engagement task. For F0, we found a main effect of information status, with higher F0 for new and contrast conditions, and an effect of engagement, with overall higher F0 in the high-engagement task. Lastly, we found significantly higher intensity for the high-engagement task.

These data suggest that speaker engagement matters in the production of prominence and that speakers may provide more reliable cues to prosodic context when they are engaged in a conversation.

I remember connectionism

Mark S. Seidenberg (University of Wisconsin-Madison)
seidenberg@wisc.edu

Connectionist ideas and computational models have been around for over 25 years now. What is the future of this type of work in a world in which there are other attractive frameworks (e.g., Bayesian) and methods (e.g., neuroimaging, big data)? Believing that what's past is prologue, I'll examine this question by looking at research in three areas: Catastrophic interference, language acquisition, and the brain bases of behavior. In each case, phenomena were discovered that were thought (and are still widely believed) to establish fatal flaws in the connectionist approach. I'll look at what happened in each area, which was the same thing. Whereas the initial concerns were overstated (an understatement), there turn out to be very interesting *related* issues of continuing interest, which any viable theoretical framework has to address.

Grammatical constraints on phonological encoding in speech production

Jordana Heller & Matt Goldrick (Northwestern University)
jordana@u.northwestern.edu

Models of retrieval and encoding of phonological information during speech production have frequently focused on isolated words. To better understand how grammatical encoding influences such processes, we examine how grammatical class constraints influence the activation of phonological neighbors (words differing by the addition, substitution or deletion of one phoneme, e.g., *moon, soon* for target TUNE). Specifically, we compare how neighbors that share the grammatical category of a target word influence picture naming latencies when grammatical constraints are strong (in a sentence context) vs. weak (bare noun naming [1]). Our results show that within-category neighbors significantly influence phonological encoding specifically within sentence contexts, suggesting that grammatical encoding constrains activation during phonological processing.

Many theories of lexical access incorporate feedback between lexical and phonological levels. This activates lexical neighbors (e.g., TUNE activates /un/, which in turn activates *moon* and *soon*), allowing these non-target representations to influence processing. Studies of bare noun naming have shown that the number of neighbors (*neighborhood density*) can facilitate target processing [2]. However, under other circumstances, activation of phonologically related words may inhibit target processing [3]. We examine whether density effects (either facilitatory or inhibitory) are modulated by grammatical category constraints, specifically in production contexts where grammatical constraints strongly influence processing. If, under these processing conditions, the activation of neighbors is subject to grammatical constraints, retrieval of the target word in a sentence should be influenced only by neighbors that share its grammatical category (e.g., for TUNE, neighbors like *moon* should influence retrieval, whereas neighbors like *soon* should not). This predicts that within-category neighborhood density should influence target retrieval. Conversely, if phonological processing is not influenced by grammatical constraints, all neighbors (regardless of category) should influence processing In this case, within-category density should not exert any influence on processing over and above the effect of overall neighborhood density.

To examine these contrasting accounts, we constructed two sets of nouns matched for overall neighborhood density, but differing in number of within-category neighbors. 64 participants named 22 pictures representing these target nouns both in bare picture naming and in sentence contexts. In the latter, participants read aloud a non-semantically predictive sentence frame (presented word-by-word) and then named a picture to complete the sentence. Log-transformed naming latencies were analyzed with respect to number of within-category neighbors, contrast-coded condition, block number, and their interactions using a mixed-effects regression model including the maximal supported random effects structure. Significance values were ascertained by model comparison. As shown in the figure below, within-category neighbors lead to slower latencies in the sentence condition, but not in bare naming (supported by a significant interaction of naming condition and within-category density; $\chi2(1) = 4.8$, $p < 0.05$). These data suggest that in production contexts with strong grammatical constraints, phonological activation is constrained by grammatical encoding processes. This supports theories integrating phonological and grammatical encoding processes in speech production.
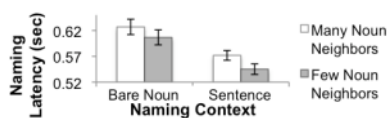
References [1] Dell, G. S., Oppenheim, G. M., & Kittredge, A. K. (2008). Saying the right word at the right time Syntagmatic and paradigmatic interference in sentence production. *Language and Cognitive Processes, 23,* 583-608. [2] Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 735-747. [3] Jaeger, T. F., Furth, K. & Hilliard, C. (2012). Incremental phonological encoding during unscripted sentence production. *Front. Psychology,* doi: 10.3389/fpsyg.2012.00481.

More than words: The effect of multi-word frequency and constituency on phonetic duration

Inbal Arnon (University of Haifa) & Uriel Cohen Priva (Brown University)
Inbal.arnon@gmail.com

Recent evidence shows that language users are sensitive to the distributional properties of multi-word sequences [e.g.,1, 2]. Such findings undermine the traditional distinction between stored and computed forms and call for processing models that can represent larger chains of relations. Previous findings show that production is sensitive to multi-word frequency: children show phonetic reduction for higher frequency phrases [3] while adults are faster to start producing them [4], and more likely to omit material in frequent contexts [5]. Here, we go beyond previous findings by asking if the effect of multi-word frequency on phonetic duration is modulated by syntactic constituency. We show that (a) that multi-word statistics affect phonetic duration in adults in spontaneous speech, and (b) that the effect is not modulated by constituency: duration is shorter for higher frequency sequences within and across syntactic boundaries. We looked at the effect of multiword frequencies and constituency on phonetic duration in spontaneous speech in two studies (using the time-aligned and parsed portions of the Switchboard corpus). In the first study, we contrasted two three-word syntactic structures: (a) subject-auxiliary-verb sequences (e.g., *everybody was trying*) – which are not constituents (Marantz 1981), and (b) verb-determiner-complement sequences (e.g., *saw the boy*) -- which are constituents. If constituency plays a significant part in the retrieval and production of multiword sequences, we should see a smaller (or no) effect of frequency on the duration of trigrams in non-constituent sequences compared to trigrams in constituent sequences.

We removed sequences more of less than 2.5sds from the mean duration (N=2515 for constituents, N=2289 for non-constituent). We controlled for various phonetic and phonological variables (mean log(word duration in the corpus), log(speaker-speech-rate), log(number-of-syllables), end-of-phrase) as well as for unigram and bigram frequencies. Trigram frequency emerged as a significant predictor regardless of constituency: phonetic duration was reduced in higher frequency sequences for both constituents (β=-0.004, $p < 0.01$) and non-constituents (β=-0.007, $p < 0.05$). In the second study – to ensure our findings were not limited to the constructions used in Study 1 - we examined duration in multiple three-word constituent and non-constituent structures in post-verbal positions. Constituents were any single post-verbal phrase: VP[V XP], and non-constituents were any two post-verbal phrases: VP[V XP YP] (e.g., *pruning once a year* vs. *doing this to you*). Using the same procedure as in Study 1 (N=5280), we again found that phonetic duration was reduced in higher frequency sequences regardless of constituency (constituents: β=-0.003, $p < 0.05$, non-constituents: β=-0.005, $p < 0.05$, and no interaction, $p > 0.6$).

This is, to our knowledge, the first study to systematically investigate the interaction of multi-word statistics and constituency. The findings show that sensitivity to surface distributions (multi-word frequency) is not affected by the higher order properties of the sequence in question (constituency). While the reported findings are limited to production, and may be driven by prosodic, rather than syntactic factors, they highlight speakers' sensitivity to multi-word frequency and open up a novel set of questions about the interaction between surface distributions and higher order properties, and the resulting need - or lack thereof - to incorporate higher order properties into processing models.

References. [1] Arnon, I. & Snider, N. (2010). *Journal of Memory and Language*, 62, 67-82. [2] Tremblay, A., Derwing, B., Libben, G., & Westbury, G. (2011), *Language Learning*, 61, 569-613. [3] Bannard, C., & Matthews, D. (2008). *Psychological Science*, 19, 241-248. [4] Jannsen, N., & Barber, H. A. (2012). *PLoS ONE 7(3): e33202. doi:10.1371/journal.pone.0033202.* [5] Levy, R., & Jaeger, T. F. (2007). In B. Schlokopf, J. Platt, & T. Hoffman (Eds.). *Advances in neural information processing systems* (NIPS) (Vol. 19, pp. 849–856). Cambridge, MA: MIT Press.

Syntactic adaptation: Converging on the statistics of the linguistic environment

Alex B. Fine (University of Rochester), Thomas A. Farmer (University of Iowa) & T. Florian Jaeger (University of Rochester)
abfine@gmail.com

Language is variable in that language is *used differently* in different situations (e.g., by different speakers). This holds at multiple levels of linguistic representation, including syntax [1-2]. Given evidence suggesting that language comprehension is sensitive to the statistics of previous linguistic experience [3], and the claim that experience-based parsing leads to efficient processing [4], variability raises a fundamental question for models of parsing: how do we benefit from experience-based processing when the statistics of individual situations vary? We propose that comprehenders adapt to novel linguistic environments such that their expectations *converge towards the statistics of the environment*. This proposal extends and offers a broad conceptualization of work on syntactic priming [5-7]. We test our proposal with three self-paced reading experiments. The experiments manipulated the relative proportion of temporarily ambiguous relative clauses (RCs) and matrix verb sentences (MVs; cf. (1)-(4); mixed with fillers containing no such ambiguities). In all experiments, the "ambiguity effect"—the difference in reading times at the underlined regions in ambiguous vs. unambiguous RCs—diminished as a function of the relative probability of RCs *in the experiment* compared to *everyday language use*. Crucially, the ambiguity effect for MVs (the a priori more frequent structure) either remained unchanged (when their relative probability in the experiment was comparable to real life experience) or *in*creased (if their relative probability in the experiment was smaller than in real life). The latter is illustrated for Exp. 3, which manipulated the amount of RC exposure in a blocked, between-subjects design (see table). As RCs are assigned *higher* subjective probabilities, MVs should be assigned *lower* subjective probabilities, and the ambiguity effect for MVs (RTs in 3 vs. 4) should *increase* as more RCs are observed and as the ambiguity effect for RCs *decreases*. The figure plots the ambiguity effect across blocks for both groups. The ambiguity effect for RCs is diminished in Block 2 for the RC-First but not the Filler-First group. Crucially, the ambiguity effect for MVs in Block 3 is greater for the RC-First than the Filler-First group: the more RCs subjects see, the *more* they expect RCs, and the *less* they expect MVs. In short, we find that the *local* environment-specific statistics of syntactic structures can both speed up and slow down processing. Our findings are *in*compatible with transient activation [7] and episodic memory [8] accounts of syntactic priming, which are insensitive to local environment-specific statistics. Instead, our findings are interpretable as cumulative syntactic priming only if (a) the magnitude of syntactic priming is sensitive to the prediction error experienced while processing the prime [9] and (b) syntactic predictions are continuously adapted to match the statistics of the current environment. Our work thus links syntactic adaptation, implicit statistical learning, and syntactic priming in one coherent framework.

Group	Block 1 (Exposure)	Block 2 (RC test)	Block 3 (MV test)
RC-First (n=40)	16 RCs (8 ambiguous)	10 RCs (5 ambiguous) + 20 fillers	10 MVs (5 ambiguous) + 15 fillers
Filler-First (n=40)	16 fillers	10 RCs (5 ambiguous) + 20 fillers	10 MVs (5 ambiguous) + 15 fillers
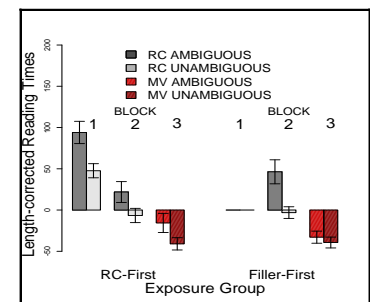
1. **Ambiguous RC:** The soldiers warned about the dangers <u>conducted the midnight</u> raid.
2. **Unambiguous RC:** The soldiers who were warned about the dangers <u>conducted the midnight</u> raid.
3. **Ambiguous MV:** The soldiers warned about the dangers <u>before the midnight</u> raid.
4. **Unambiguous MV:** The soldiers spoke about the dangers <u>before the midnight</u> raid.

References. 1. LibermanEtAl67 **2.** Tagliamonte05 **3.** MacDonaldEtAl94 **4.** SmithLevy08 **5.** AraiEtAl07 **6.** ThoathathiriSnedeker08 **7** PickeringGarrod04 **8.** KaschakGlenberg04 **9.** ChangEtAl06

Direct experience versus abstract knowledge in linguistic processing

Emily Morgan & Roger Levy (University of California, San Diego)
eimorgan@ucsd.edu

Abstract linguistic knowledge uncontroversially plays an integral role in the processing of novel expressions, producing differential difficulty according to an expression's grammatical and semantic properties. It remains a major outstanding question, however, what role this abstract linguistic knowledge plays in determining processing difficulty for expressions that are *not* novel—those with which the speaker has had direct experience. Arnon & Snider (2010), for example, found processing advantages for multi-word expressions which could not be accounted for by frequencies of individual words or substrings. Likewise, many *binomial expressions* (of the form "X and Y"; McDonald et al., 1993) have a preferred word order (e.g. "bride and groom" vs. "groom and bride"), and frequently attested binomials—those which native speakers almost certainly have direct exposure to—are read faster in their more frequent order (Siyanova-Chanturia et al., 2011). It has been shown that ordering preferences of attested binomials can be predicted with high accuracy from a small number of semantic, phonological, and lexical factors (Benor & Levy, 2006). Do the results seen in these studies derive strictly from the underlying grammatical and semantic factors giving rise to the differences in expression frequency, or are they mediated by native speakers' frequency of direct exposure to the expressions?

To answer this question, we developed a logistic regression model based on six of the strongest factors from Benor & Levy to predict direction and strength of binomial ordering preferences. We then constructed 42 novel binomials (unattested in the 500-billion-word Google books corpus, and thus likely to be novel for most speakers) for which our model predicts strong ordering preferences. In an offline forced-choice ordering-preference study using these novel binomials and 42 attested binomials from Siyanova-Chanturia et al., we found that our model predicted native-speaker ordering preference well for both novel and attested binomials, but that for any given strength of ordering preference in our model, speaker preferences were more extreme for attested binomials than for novel binomials.

We then conducted a self-paced reading study (400 participants) in which these binomials were embedded in sentence context and presented either in preferred or non-preferred order (interspersed with filler sentences). Both attested and novel binomials were read significantly faster in preferred than in non-preferred order according to our model's dichotomous ordering preference predictions. However, in a linear mixed-effects regression model, our model's quantitative prediction for strength of ordering preference was a significant predictor of reading times only for novel binomials, not for attested binomials. In a further regression model combining both our model's quantitative prediction and the true dichotomous ordering preference (according to Siyanova-Chanturia et al.), only the true preference was a significant predictor of reading times for attested binomials.

We conclude that at least some abstract linguistic constraints determining binomial ordering preferences are active in the online comprehension of novel binomials, but that for attested binomials, exposure to specific multi-word expressions overwhelms effects of underlying constraints. These results lend support to models of language structure and form in which both productivity and reuse play a role (O'Donnell et al., 2011; Johnson et al., 2007). Abstract knowledge is essential in the processing of novel expressions, but in the processing of expressions that have been encountered before, the full apparatus of linguistic productivity may not be necessary, as these expressions are stored and may be reused directly.

References. Arnon, I., & Snider, N. (2010). *JML*, *62*(1). Benor, S., & Levy, R. (2006). *Language*, *82*(2). Johnson, M., Griffiths, T. L., & Goldwater, S. (2007). *NIPS*, *19*. McDonald, J. L., Bock, K., & Kelly, M. H. (1993). *Cognitive Psychology*. O Donnell, T., Snedeker, J., Tenenbaum, J. B., & Goodman, N. (2011). *Proc CogSci*. Siyanova-Chanturia, A., Conklin, K., & van Heuven, W. J. B. (2011). *JEP:LMC, 37*(3).

A rational inference approach to aphasic language comprehension

Edward Gibson (MIT), Chaleece Sandberg (Boston University), Ev Fedorenko (MIT), Swathi Kiran (Boston University)
egibson@mit.edu

Broca's aphasics have long been known to preferentially use plausibility information when it conflicts with syntactic information (e.g., Caramazza & Zurif, 1976). For example, an implausible sentence like "*The girl was kicked by the ball*" is often interpreted by Broca's aphasics according to semantic cues ("*the girl*" – agent; "*the ball*" – patient). The same is not true for healthy participants (e.g., MacWhinney & Bates, 1989; Ferreira, 2003). We propose that this property of aphasic comprehension follows from a Bayesian approach to sentence comprehension (Shannon, 1949; Bates et al., 1991; Levy et al., 2009). In a linguistic exchange, the comprehender's guess of what the speaker intended can be formalized as the probability of the speaker's intended meaning (s_i) given the perceptual input (s_p): $P(s_i \mid s_p)$. By Bayes' rule, this is achieved by multiplying the prior (i.e., what is likely to be said), $P(s_i)$, with the likelihood that a noise process would generate s_p from s_i, $P(s_i \rightarrow s_p)$. We hypothesize that because aphasic individuals are aware that their language comprehension system is damaged, they assume greater probability of noise than healthy comprehenders. This assumption leads to a greater reliance on the prior (i.e., plausibility of the utterance).

This hypothesis was evaluated in a paradigm where participants are asked to act out the meaning of a spoken sentence using iconic cards on sticks. Three groups were tested: 8 aphasic patients with documented sentence comprehension deficits; 10 age-matched controls; 12 young controls. Two English alternations were examined: the active/passive which requires at least two insertions or deletions to get from an implausible version to its plausible counterpart (1,2); and the double-object (DO) / prepositional-phrase object (PO) which requires only one insertion / deletion to get from an implausible version to its plausible counterpart (3,4).

(1) Active plaus: The girl kicked the ball. → Passive implaus: The girl <u>was</u> kicked <u>by</u> the ball.
(2) Passive plaus: The ball <u>was</u> kicked <u>by</u> the girl. → Active implaus: The ball kicked the girl.
(3) DO plaus: The boy gave the girl the book. → PO implaus: The boy gave the girl <u>to</u> the book.
(4) PO plaus: The boy gave the book <u>to</u> the girl. → DO implaus: The boy gave the book the girl.

In previous work, Gibson, Bergen & Piantadosi (2012) demonstrated that healthy comprehenders rely more on the prior for an implausible sentence when fewer edits are required to arrive at a plausible alternative. Thus, they rely more on the prior in interpreting implausible DO/PO structures than implausible actives/passives. Furthermore, they rely more on the prior for *deletions from* a plausible sentence (as in (4)) than for *insertions to* a plausible sentence (as in (3)), consistent with the Bayesian size principle (Tenenbaum & Xu, 2000; McKay, 2003).

Results. *First*, aphasics reliably use plausibility in both the active/passive and DO/PO alternations. In contrast, controls rely on plausibility only in the DO/PO alternation (not pictured; replicated from Gibson et al., 2012). *Second*, aphasics, like controls, rely on plausibility more in deletions from plausible sentences than insertions to plausible sentences for the DO/PO structures. *Third*, and most critically, aphasics, like controls, rely on plausibility *more* in the DO/PO than in the active/passive alternation, as evidenced by a reliable interaction between plausibility and construction (multiple-edit = active-passive vs. single-edit = DO/PO).

These data provide strong evidence for the noisy-channel communication-based proposal: Aphasics perform rational inference in sentence comprehension just as healthy controls do, but in maximizing $P(s_i \mid s_p)$, they assume a higher noise rate, $P(s_i \rightarrow s_p)$, than healthy comprehenders, with the consequent greater reliance on their semantic priors, especially in the minor-edit alternations.

Comprehension and acquisition of contrastive prosody: Rational inference helps adults and children cope with noisy input

Chigusa Kurumada (Stanford University), Meredith Brown & Michael K. Tanenhaus (University of Rochester)
kurumada@stanford.edu

Listeners can use pitch accent and intonation to derive pragmatic meanings in spoken language (e.g., Pierrehumbert & Hirschberg, 1990; Ito & Speer, 2008; Dennison & Schafer, 2010). However, the acoustic features of relevant representations are notoriously variable (e.g., H* vs. L+H*, Watson, Gunlogson, & Tanenhaus, 2008; Arnold & Watson, submitted). We present novel evidence that listeners overcome the challenge of massive variability and noise in the prosodic signal through rational inference, as has been attested in phoneme perception (e.g., Kraljic, Brennan, & Samuel, 2008).   Moreover, we demonstrate that rational inference allows children to comprehend contrastive prosody much earlier than previously reported.

Experiment 1 (*n*=56, 24 trials) tested whether form-based inferences modulate prosodic interpretation in adult listeners. We used the construction "It looks like an X" pronounced either with 1) nuclear H* on the final noun (**Noun-focus prosody**), indicating that the referent is likely to be X, or 2) contrastive L+H* on "looks" (**Verb-focus prosody**), implying otherwise (e.g., *It LOOKS like a butterfly but it is not one*). The task was a 2AFC decision between a picture of the mentioned referent X (e.g., a butterfly) and a picture that visually resembled X (e.g., a moth). Participants were randomly assigned to one of two between-subjects conditions. In the **Prosody-only condition**, participants heard "It looks like an X" with either Noun-focus or Verb-focus prosody (12 each). In the **Prosody+Form condition**, participants heard these two prosodic patterns (8 each) as well as 8 instances of "It's an X" (e.g., "It's a BUTTERFLY!") with nuclear accent on the final noun. **[Results]** Although Verb-focus prosody elicited more "not X" responses in both conditions (*p*<.001), participants were also more likely overall to respond "not X" in the Prosody+Form condition than in the Prosody-only condition (*p*<.05), presumably on the pragmatic expectation that the speaker would have said "It's an X" if she had meant so.

In **Experiment 2** (*n*=36, 16 trials), 4-year-olds were tested in a similar 2AFC picture-selection paradigm. They selected a likely referent based on a puppet's utterances. In the **Prosody-only condition**, as in Exp1, the puppet always said "It looks like an X", encoding the two meanings (i.e., the referent is X or not) using Noun-focus vs. Verb-focus prosody. In the **Form-only condition**, the puppet instead used two constructions: "It's an X" vs. "It looks like an X" (with Noun-focus prosody). In the **Prosody+Form condition**, the puppet used both constructions and marked "It LOOKS like an x…" with Verb-focus prosody. **[Results]** 4-year-olds did not distinguish between the two prosodic patterns in the Prosody-only condition. However, when the input also included a more reliable formal contrast (Prosody + Form condition), children were almost at ceiling in interpreting "It LOOKS like an x…" as indicating a contrastive meaning (i.e., …***but it is not one***). This supports a view that the presence of an unambiguous expression ("It's an X") encourages children to infer that contrastive prosody signals a distinct speaker meaning.

The current results fill a major gap in the theory of pragmatic comprehension of prosody. Prosodic input is interpreted conditionally on what linguistic cues – structural as well as prosodic – are used by a specific speaker. 4-year-olds, who are otherwise apparently insensitive to the prosodic contrast between H* and L+H*, can bootstrap their knowledge through an inference based on the formal contrast. Thus we conclude that any adequate model of the pragmatic comprehension and acquisition of prosody must integrate the process of rational inference based on contextually constrained expectation.

References. Arnold, J., & Watson, D. (under review). Synthesizing meaning and processing approaches to prosody: performance matters. Dennison, H. Y., & Schafer, A. J. (2010). In Speech prosody 2010 conference. Ito, K., & Speer, S. R. (2008). JML, 58 (2), 541–573. Kraljic, T., Brennan, S. E., & Samuel, A. G. (2008). Cognition, 107 (1), 54–81. Pierrehumbert, J., & Hirschberg, J. (1990). Intentions in Communication. Watson, D. G., Tanenhaus, M. K., & Gunlogson, C. A. (2008). Cognitive Science, 32 (7), 1232-1244.

Subject relative clauses versus object relative clauses: Difference among adults and children
Yuki Hirose (The University of Tokyo) & Reiko Mazuka (RIKEN)
hirose@boz.c.u-tokyo.ac.jp

Psycholinguistic evidence on pre-nominal RCs have been controversial (see O'Grady 2011) although most findings support a processing advantage for S-RCs. With acquisition data, the situation is again mixed. For example, Japanese pre-school children actually performed better with O-RCs than S-RCs in a comprehension experiment, although the difference was apparently explained by the children's inability to use the object case marker -o (Suzuki 2011). Our study presents two identical visual world experiments with Japanese-speaking adults and first graders (6;5-7;4). Both groups (N=12 each) had sufficient knowledge and command of subject and object case markers. In additions to the RC gap type (S- vs. O-RCs), the animacy combination of the RC constituents was also manipulated to examine if the animacy effect (O-RCs become easier with an animate subject and an inanimate object: Traxler et al. 2002, 2005, Mak et al. 2002, 2006), is observed the same way in Japanese-speaking adults and children. Twelve quadruples such as a. ~ d. below were prepared with fillers (Note: The verb in the examples takes a dative direct object and it can take animate or inanimate subject.). In the experiments, participants saw eight visual objects on a screen, each the combination of two items (animal + fruit), one on top of the other (e.g., Fig.1). At the same time, they heard audio stimuli such as a. ~d. below and were asked to point at the correct referent. We calculated (i) the log odds of the looks to the Correct Target and (ii) the looks to the Incorrect Target (i.e., the reversed subject and object interpretation) against the looks to all the other objects combined, for 200ms intervals from the RC verb offset. The Anim. Subj. + Inanim Obj. versions (a. & b.) and Inanim. Subj.+ Anim. Obj. versions (c. & d.) were compared separately. For RCs with Anim. Subj. and Inanim. Obj. (a. and b.), adults and children exhibited opposite patterns and the effect of the RC-type mainly emerged in the looks to the Incorrect Target. Adults produced more looks to the Incorrect Target when they heard O-RC (b.) than S-RC (a.) in the 400-600ms interval following the RC verb offset and onward ($p < .05$). Adults looked more at the Correct Target, when it was referred to by a S-RC (a.) compared with an O-RC (b.) but not until in later intervals (1200-1400ms-). Children, in contrast, produced more looks at the Correct Target when it was referred to by O-RCs and more looks to the Incorrect Target when it was referred to by S-RCs, though the effect showed up much later than in adults ($p<0.5$ for both comparisons, in intervals 2200ms after the RC verb offset and onwards). For RCs with Inanim. Subj. and Anim. Obj. (c. and d.), adults looked more on the Incorrect Target when referred to by an O-RC than a S-RC (800-1000ms interval and onwards, $p<.05$) and so did children at a later interval (1400-1600ms, $p<.05$). No significant difference was found for looks to the Correct Target in either group for c. and d. In short, we found a mirroring pattern where S-RCs are easier for adults while O-RCs are easier for children, except that O-RCs with Inanim. Subj. and Anim. Obj led to more erroneous looks in both children and adults. The animacy effect reported in the literature was found in both adults and children. To account for the otherwise opposite pattern between adults and children, we will provide discussions from the perspective of (i) the possible difference between adults and children in the efficiency in using the case marker information and (ii) reconsideration of the linear distance account in RC processing whose effect appears to be absent in Japanese adults possibly for independent reasons.

a. S-rel (Anim. Subj. + Inanim. Obj.) :
   e   ringo-ni   notteru   ahiru-wa   doko?
      apple-DAT be on top duck-TOP where
   "Where's the duck which is on top of the apple?"
b. O-rel (Anim. Subj. + Inanim. Obj.) :
   ahiru-ga e   notteru   ringo-wa   doko?
    duck-NOM   be on top apple-TOP where
   "Where's the apple which the duck is on top?"
c. S-rel (Inanim. Subj. + Anim. Obj.) :
   e ahiru-ni   notteru   ringo-wa   doko?
      duck-DAT   be on top apple-TOP where
   "Where's the apple which is on top of the duck?"
d. O-rel (Inanim. Subj. + Anim. Obj.) :
   ringo-ga e   notteru   ahiru-wa   doko?
   apple-NOM   be on top duck-TOP where
   "Where's the duck which the apple is on top?"

The Correct Target for a. and b./ The Incorrect Target for c. and d.

The Correct Target for c. and d./ The Incorrect Target for a. and b.

Figure 1. Sample visual scene

Poster Session 3
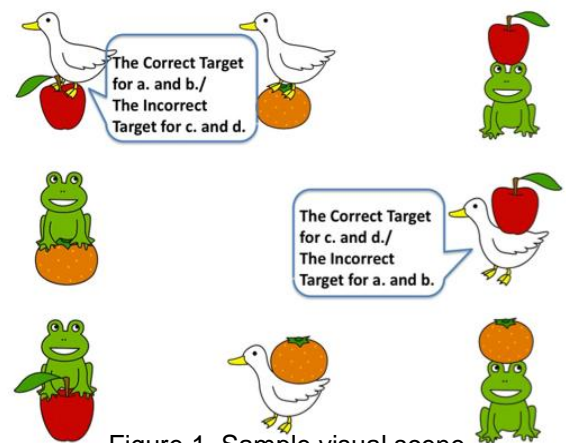
Levels of syntactic representation in bilingualism

Guadalupe de los Santos & Julie Boland (University of Michigan)
gdelossa@umich.edu

We conducted the present study to see how syntactic information is shared across languages in skilled bilinguals. We wished to distinguish between the one-level word-order account (Loebell & Bock, 2003) and two-level accounts that argue for a word-order level and an abstract functional level (Shin & Christianson, 2009; Hsieh, 2010). The two-level account argues for a higher level of shared-representation, suggesting that languages are linked on a functional level, not simply on the surface.

Using a sentence production paradigm adapted from the Loebell and Bock study, we tested within syntactic structure, using the *se*-passive in Spanish as a prime for English passives. This is in contrast to the original Loebell and Bock study which compared priming in datives and passives for priming differences in word order. By testing within syntactic structure, we reduce the variability of the strength of priming effects; an issue since priming effects are subtle and dative structures have been shown to have a stronger priming effects than passive structures in English (Bock 1986). By using a picture description task, we are looking at sentence production rather than comprehension, which has been more consistently sensitive across studies testing syntactic priming (Tooley & Traxler, 2010). Previous studies which support the two-level account either used priming in comprehension (Hseih, 2010) or a sentence-recall task (Shin & Christianson, 2009).

During our study, Spanish-English bilinguals repeated an auditory passive priming sentence in Spanish and then described an unrelated, transitive image in English (e.g., a shark biting a person). The passive priming sentence would have either the same word order as English (NP+V) or an alternating structure (V+NP) which is also valid in Spanish. Samples of the stimuli are below. The sentence produced by the participant was then coded as passive, transitive, or other (intransitive/descriptive/reflexive).

If the one-level word-order account is correct, we would expect to see priming only when the prime had the same word order as an English passive, NP+V. The two-level account would be supported if the same amount of priming occurs across both prime sentence word-orders. (Both theories predict priming when the word-order is the same, but only the two-level account predicts priming when the word order is different.)

We found the same amount of passives produced in English regardless of the word order of the prime. This result directly contradicts Loebell and Bock's (2003) claim that cross-linguistic priming is limited to constructions that have the same word order in both languages. Our results support the two-level account of syntactic representation. We believe that the difference in results is because of testing within syntactic structure rather than between syntactic structures.

    'Se perdieron las llaves.'   'Las llaves se perdieron.'
    PRN lost   the keys   the keys PRN lost
    The keys were lost.   *The keys were lost.*

    PRN- Pronoun; without pronoun 'se' sentences are active transitive sentences (Zagona, 2002).

Preschool-aged children process words and sentences talker-contingently

Sarah C. Creel (University of California, San Diego)
creel@cogsci.ucsd.edu

Many approaches to language processing relegate talker information to the role of extralinguistic context, suggesting it is processed by different cognitive (McLennan & Luce, 2005) or neural (Belin et al., 2000) mechanisms than language. Yet, other studies hint that young infants store words talker-specifically (at 7.5 months, but not 10.5; Houston & Jusczyk, 2000), and that adults show sensitivity to talker-specific word information (Goldinger, 1996). It is not clear if infants' acoustically-specific representations are homologous to adults' talker sensitivity. The current study tested talker specificity at an intermediate age—3-5 years.

In Exp. 1, 32 children learned two word pairs as labels for novel cartoon creatures: a *same-talker* pair (e.g. a male spoke both the words *geeb* and *geege*), and a *different-talker pair* (male *marv*, female *mard*). One pair was learned in 16 exposure trials (e.g. "That's a geeb!"), followed by 16 eye-tracked test trials ("Where's the geeb?"; then the other pair was trained and tested (order counterbalanced). If children do not store talker information with words, they should not recognize different-talker words any faster than same-talker words. If they *do* store talker information, they should look faster to referents of different-talker words than same-talker words. Children fixated referents faster in the different-talker condition (Fig. 1): the target advantage score (target minus other-picture looks) was greater in different-talker trials than same-talker trials from 600-700 ms (*p*=.03), 700-800 ms (*p*=.006), and 800-900 ms (*p*=.03). Moreover, target advantage exceeded chance (0) sooner for different-talker trials (400-500 ms, *p*=.04) than for same-talker trials (900-1000, *p*=.02). Thus, 3-5-year-olds show adultlike talker-specific word recognition. Accuracy did not differ (different, 72% vs. same, 66%, *p*=.24), implying that talker information was implicit, or at least not strategic.
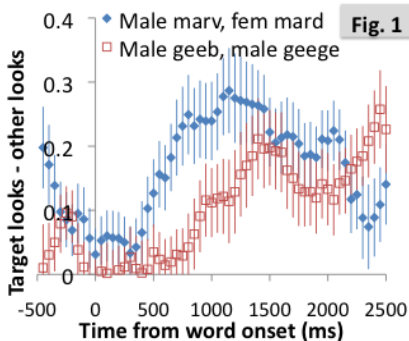
Figure 1. Looks in Experiment 1, ± standard error.
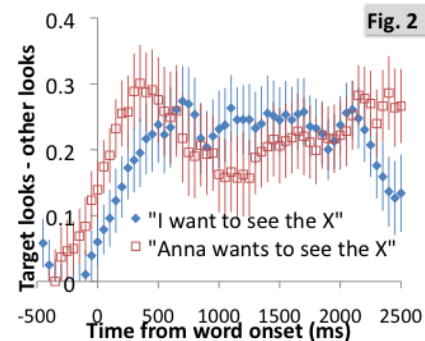
Figure 2. Looks in Experiment 2, ± standard error.

Were children storing talker acoustics, or associating *talker identity* and referents—a more semantic/pragmatic use of talker—or both? Exp. 2 trained 32 children on different-talker pairs. During training, talkers expressed strong preference for the picture they labeled ("Yay, the geeb is my favorite one!") Children were then tested on 1st-person trials ("I want to see the geeb") and 3rd-person trials ("Anna wants to see the geege"). The "wanter" always matched the wanter at training, but the voice only matched the training voice on 1st-person trials. If children stored only *acoustics*, then 3rd-person trials should show slow looks and lower accuracy. However, if children encoded *who liked what*, then 1st- and 3rd-person trials should be similarly fast and accurate. Looks to target exceeded chance before word onset (Fig. 2) in both trial types (-100 to 0 ms, *p*=.003), with equivalent accuracy (1st: 71%; 3rd: 73%; *p*=.54). This suggests that children readily associate novel referents with talkers, using voice to decide whose favorite referent to fixate.

In sum, children interpret words and sentences talker-contingently. Ongoing work examines whether effects are primarily acoustic, semantic, or both. Talker sensitivity in 3-5-year-olds suggests developmental continuity in talker-specificity effects. Results cast doubt on theories that talker information is extralinguistic, suggesting instead that talker information pervades language processing from the earliest moments of language input.

Baseball bats & butterflies: Context effects of on pragmatic inferencing in adults and children
Yi Ting Huang & Alix Kowalski (University of Maryland College Park)
ythuang1@umd.edu

Speakers often convey more than they say, thus listeners must generate inferences that allow them to access the intended meaning. Previous research has recruited inferences like scalar implicatures to shed light on the semantics-pragmatics interface (Huang & Snedeker, 2009ab). This work has shown that for lexicalized scales, both adults and children initially interpret the semantics of utterances ("*some*" = POSSIBLY ALL) and only later restrict meaning via inference (NOT ALL). This suggests that developmental delays in acquiring implicatures in off-line tasks mirror those seen in real-time comprehension. However surprisingly, recent work suggests that young children generate implicatures for ad-hoc scales (Stiller et al., 2011). Despite profound difficulties with lexicalized scales, they easily restrict interpretation of sentences like (1) to subsets ($girl_1$).

(1) Click on the girl that has <u>baseball</u> bats   ($girl_1$: baseball bats, $girl_2$: baseball bats + butterflies)

This precocity may reflect distinctions in the relationship between language and context across these scales. While lexicalized scales draw contrast through linguistic alternatives (*some*, *all*), ad-hoc scales do so through discourse context. Critically, the exact nature of these context effects remains unknown. One possibility is that context leads listeners to encode referents informatively, even before the onset of linguistic input ($girl_1$ = "*girl with baseball bats*", $girl_2$ = "*girl with butterflies*"). Alternatively, it may be that context effects only influence interpretation after the onset of the linguistic trigger ("*baseball*").

The current study presented adults and 5-year-olds with stories about objects given to four characters. Afterwards participants heard directions like (1) while their eye-movements were measured to subsets ($girl_1$) and supersets ($girl_2$). In these critical trials, reference restriction requires a pragmatic inference ("*baseball bats*" = NOT BUTTERFLIES). Critically, if discourse context promotes informative encoding of referents before (1), then an early preference for the subset should emerge at the linguistic trigger. If, however, context is only enlisted after the trigger, then both subsets and supersets will initially be considered. Control trials used the same displays to ask for "*only baseball bats*" and "*butterflies.*" Since these referents are semantically disambiguated, preferences for correct targets here are expected to emerge early.

Analysis of character selections in the control trials indicated overwhelming preferences for correct targets in both adults and children (>85%). Furthermore, analysis of fixations revealed early preferences for supersets after the onset of the linguistic trigger. However, while adults also demonstrated an early preference for subsets, children's fixations suggested difficulties with the focus operator ('*only*'). Overall patterns in these control trials indicate rapid semantic disambiguation and parallels between off-line and on-line interpretations. In contrast, actions in the critical trials unanimously favored the subsets (100%), but fixations following the trigger were equivocal. Target preferences did not emerge until 658ms after the trigger in adults and 858ms in children. This demonstrates that context facilitated inferencing in ad-hoc scales, but its effects do not emerge until after the linguistic trigger. These delays in real-time interpretation were comparable to those found in lexicalized scales (Huang & Snedeker, 2009ab). Altogether these results suggest parallelisms between lexicalized and ad-hoc scales. In both cases, scalar expressions are initially interpreted through their semantics. However, these findings also point to possible differences. While inferences for lexicalized scales are acquired late in development, those for ad-hoc scales are acquired early. These results illustrate how a developmental perspective offers a unique window into the effects of diverse inputs (e.g., linguistic alternatives, discourse context) on the cognitive architecture of pragmatic inferencing.

TRIAL TYPE - *Linguistic trigger*	**ACTION: % target**		**FIXATION: Time from trigger**	
	Adults	**Children**	**Adults**	**Children**
Critical - *Baseball bats*	100%	100%	658 msec	858 msec
Control - *(only) Baseball bats*	93%	86%	200 msec	850 msec
Control - *Butterflies*	100%	87%	200 msec	250 msec

The abstraction of syntax by fits and starts

Nick Gruberg, Liane Wardlow & Victor Ferreira (University of California San Diego)
Contact: ngruberg@ucsd.edu

During language acquisition the child encounters utterances and communicative events from which s/he must infer the abstract structure of the language. According to usage-based theories, the language learner begins with a minimum of innate linguistic knowledge by which to structure that input. Utterances are initially processed as unitary objects, conflating phonology, lexical items, semantics and syntax. As language acquisition proceeds, the language learner begins to process utterances as fully elaborated, abstract and differentiated linguistic structures.

In the current study, we tested adults and 5 & 6 year old children in a picture-matching task, in which confederates and then subjects described the same transitive, dative, and spray-load action pictures. We were interested in how often subjects repeated the syntax of the confederate's description in their own description of the same picture later in the experiment, and the extent to which syntactic repetition was contingent on repetition of the main verb of the sentence. A usage-based account predicts that for children still acquiring language, the syntax used in a picture description will become tied to the communicative event, including the verb and the picture itself. Thus they will be more likely to repeat the syntax when describing the picture, and still more likely to do so if they repeat the main verb of the sentence. On the other hand, adults will have acquired abstract syntactic representations, which can vary independently of particular communicative events. Thus they will be less likely than children to repeat the syntax of the sentence, and repeating the main verb will have a smaller effect. Further, if experience with verbs facilitates the abstraction of syntactic structures, then for children we expect less frequent verbs to be tied more strongly to communicative events and therefore to show greater priming. This need not hold for adults, whose syntax should be uniformly abstract.

The data show that both children and adults were more likely to describe a particular picture with the same structure used by the confederate (21% and 13% respectively) as compared to the alternative structure. Children were marginally more likely to do so than adults ($p = .08$). Both adults and children were 9% more likely to repeat the structure when the verb was low frequency than when it was high frequency. When children repeated the main verb, they were 28% more likely to repeat the confederate's structure compared to the alternative, but they were 2% less likely to repeat the structure when they did not repeat the verb. When adults repeated the main verb, they were 15% more likely to repeat the structure, while they were 3% more likely to repeat the structure when they did not repeat the verb.
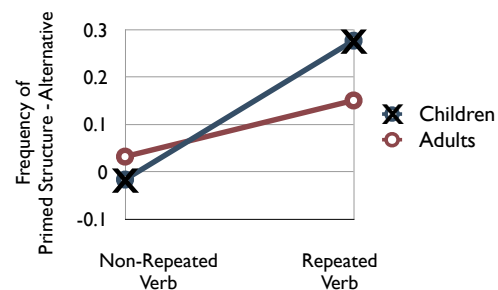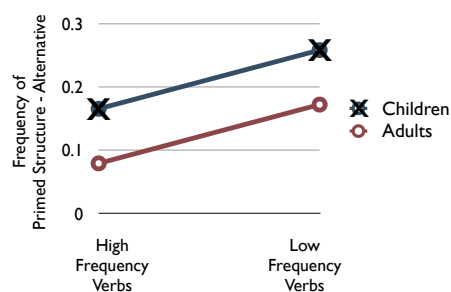
Adults and children tend to repeat the syntactic structure of picture descriptions, but this effect mostly disappears when the main verb is not repeated. This suggests that both are processing verb and syntax as somewhat associated linguistic units, an effect that seems to be more pronounced in children than adults (but see Rowland et al., 2012). Verb frequency seems to have an identical effect on children and adults, suggesting that the abstraction of syntax may not occur at a particular developmental stage, but instead may be subject to lifelong learning.

Rowland, C., Chang, F., Ambridge, B., Pine, J., & Lieven, E. (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. *Cognition* , 1-15.

The time course of filler-gap dependency processing in the developing parser

Emily Atkinson, Katherine Simeon, Akira Omaki (Johns Hopkins University)
atkinson@cogsci.jhu.edu

Much research has shown that the adult parser is biased to complete filler-gap dependencies 'actively' by associating the filler with the first thematic position (e.g., verb) [1,2]. While these studies explored the *mechanisms* of active gap filling, the *origin* of this bias is poorly understood: Is it an architectural constraint built into the parser from birth, or does it develop over time? The present visual-world study addresses this question by comparing the time course of filler-gap dependency processing in adults and 5-year-olds. We find evidence for active gap filling *only* in adults, indicating that active gap filling is a product of learning.

We adopted the visual-world eye-tracking experiment design from [3], which used a modified version of the story-based comprehension task in [2]. Each trial consisted of a) a short story with two events, and b) a comprehension question, which was either a *wh*-question (e.g., *Can you tell me what Emily was eating the cake with?*) or a *yes-no* question (e.g., *Can you tell me if Emily was eating the cake with the fork?*). The stories

Sample Display were accompanied by displays of five pictures (e.g., Story: Emily eats cake with a fork and

washes dishes with a sponge. Pictures: Emily (subject), cake (object1), fork (instrument1), dishes (object2), and sponge (instrument2); see Sample Display). Two events were included in the stories to prevent anticipations of the content of the question. The experiment presented 10 target items and 10 distractor items. Previous adult studies [2,3] found the following fixation patterns in the verb region: while there was a slight increase of fixations to the target object (cake) in the *yes-no* condition, a significantly larger increase of target fixations was observed in the *wh*-condition. This contrast reflects an active association of the *wh*-filler with the verb (i.e. the first thematic position), and constitutes visual-world evidence for the active gap filling mechanism.
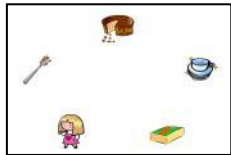
The adult eye-tracking data (N=27) replicated the previous findings [2,3] (see Fig.1). Fixations on the target object increased in the verb region in both conditions, but this increase was much larger in the *wh*-condition.
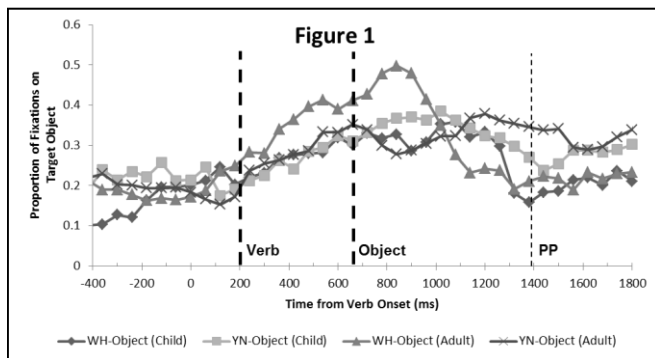
The 5-year-old eye-tracking data (N=12; mean age=5.4), however, revealed a different pattern (see Fig.1). Fixations on the target object increased during the verb region in both conditions, but the magnitude of increase was very similar in both conditions. The analysis focuses on the region in which we find effects of the verb, i.e., 200ms to 650ms after the verb onset. For the 30ms bins between 410ms and 650ms, a logit mixed model [4] with age group and question type as fixed effects and random intercepts for participants and items revealed a significant interaction of age and question type (all p<0.03). Pairwise comparisons revealed that adults' target fixations were significantly larger in the *wh*-condition (all p<0.001), while no reliable effect of question type was observed in children's target fixations (all p>0.1). High accuracy in answering questions (mean=95%) suggests that children's divergence is not due to miscomprehension of the questions.

These eye movement data indicate that, at least in the current task setting, adults actively complete filler-gap dependencies while 5-year-olds do not. This suggests the possibility that the active gap filling bias must develop over time. We will discuss a) cognitive and linguistic reasons why 5-year-olds failed to show evidence for active gap filling, as well as b) how the bias could be acquired over time by adjusting probability-based syntactic expectations through exposure to further syntactic distributional information in the input.

References: [1] Stowe, L. E. (1986). *LCP*, *1*, 227-245. [2] Sussman, R. S. & Sedivy, J. C. (2003). *LCP*, *18*(2), 143-163. [3] Omaki, A., Trock, A., Wagers, M., Lidz, J., Phillips, C. (2009). Poster presented at CUNY 2009, Davis, CA. [4] Jaeger, F. (2008). *J Memory & Language*, *59*, 434-446.

Can subset principles guide L2-Chinese learners to unlearn inverse scope? Evidence from self-paced reading

Liyuan Li & Fuyun Wu (Shanghai International Studies University)
fywu@shisu.edu.cn

Scope interpretation is known to differ in English and Chinese [1,3,4]. For sentences such as (1) containing negation **meiyou** 'not' and universal quantifier **mei** 'every', English allows both surface- and inverse-scope readings [1], whereas Chinese allows surface-scope only [2,3].This imposes challenges to second language (L2 learners whose native language is English, because they do not have negative evidence or explicit instructions to unlearn the inverse-scope reading. The Subset Principle hypothesis [4] predicts that Universal Grammar (UG guides L2 learners to acquire the unmarked surface-scope reading. But the Transfer hypothesis [4] predicts that L2 learners as influenced by their L1 have difficulties in removing the inverse-scope reading.

To further investigate these predictions, we tested Chinese native speakers (N=30) and L2-Chinese learners (N=29), using a moving-window self-paced reading task combined with truth value judgment. Participants read a short story (2) that supported either surface-scope reading (2a) or inverse-scope reading (disallowed in Chinese) (2b). Then a critical statement (1) was presented word by word. Participants judged whether it was true or false based on their understanding of the story.

Results showed that similar to Chinese native speakers, adult advanced L2 learners i) were highly accurate in rejecting the inverse scope as 'False' (91.67%) and in accepting the surface scope as 'True' (88.69%), and ii) were significantly faster in judging surface scopes than in judging inverse scopes($t = 6.42$, $p< 0.0001$). These results suggest that consistent with the Subset Principle hypothesis in L2 acquisition, L2-Chinese learners can unlearn the inverse scope reading disallowed in Chinese. This supports that UG is accessible to adult L2 learners.

But surprisingly for both groups, across most regions of critical sentences (beginning from **meiyou** 'not' for L1 and <u>dou</u> 'all' for L2 in ex.(1)), reading times were significantly shorter in the inverse-scope condition than the surface-scope condition (t's > 2.7, p's < .05). This may be due to similarity-induced interference in retrieving events during on-line processing [5,6]. Specifically, up until the word "every classmate/ all / not…,"participants couldn't decide which exactly (table-cleaning vs. kitchen-arranging) was not done by everyone in the surface-scope condition. But they didn't have to resolve event details in the reverse-scope condition. In addition, we speculate that 'dou' in the critical stimuli might also have cued participants to reset parameter values.

(1) **Mei**ge/ tongxue/ <u>dou</u>/ **meiyou**/ ca/ zhuozi/, tamen/ gengyuan/ zhengli/ chufang.
  '**Every** classmate did **not** clean the table; they were more willing to arrange the kitchen.'
(2) Xiaohong was cleaning the room at home. Three classmates visited her. They offered to help. She asked them to clean the kitchen table. The table looked a mess.
  a. *Surface-scope*: Students told Xiaohong that they were more willing to clean the kitchen.
  b. *Inverse-scope*: One student immediately started table-cleaning. The other two told Xiaohong that they were more willing to arrange the kitchen.

References. [1] May, R. (1977). The Grammar of Quantification. Cambridge, MA . [2] Aoun, J., & Li, Y.H. (1989) Scope and constituency. Linguistic Inquiry, 20, 141–172. [3] Huang, C.T. (1982). Logical relations in Chinese and the theory of grammar. Ph.D. dissertation. MIT.   [4] White, L. (1989). Universal Grammar and Second Language Acquisition. John Benjamins. [5] Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. JEP, 27(6), 1411-1423. [6] Gordon, P. C., Hendrick, R., & Johnson, M (2004). Effects of noun phrase type on sentence complexity. JML 51(1), 97-114.

L1/L2 differences in processing verbal vs. adjectival short passive constructions

Damon Tutunjian & Marianne Gullberg (Lund University)
damon.tutunjian@humlab.lu.se

Previous research suggests that "implicit" agent arguments are activated into the discourse during online comprehension of short passive constructions (Mauner & Koenig, 1999). Yet some researchers (e.g., Kratzer, 1996) have claimed that there exists a distinction between stative/adjectival and eventive/verbal passives, and that only the latter is able to license and project an agent. In addition, the two forms are superficially similar in English, which poses a potential problem for L2 learners of English when their L1s use distinct forms to mark whether a passive construction is eventive or stative. One such case is Swedish, where *vara* (*to be*) marks the stative and other forms mark the eventive usage.

Using a word-by-word reading paradigm, we tested monolingual English speakers and advanced Swedish L2 learners of English for differences in agent activation between short adjectival (1) and verbal (2) passive constructions. We hypothesized that for native English speakers, verbal passive participles would activate and make available an agent in the discourse, thus facilitating processing at an agent-dependent, rationale clause (*to keep…*), whereas adjectival passives would not. For the Swedish L2 learners of English, we explored a possible L1/L2 transfer effect (e.g., Frenck-Mestre & Pynte, 1997) whereby superficial similarities between Swedish *vara* and English *to be* would support adjectival activation and preclude agent activation, causing processing difficulty for both conditions (1) and (2).

Two processing constraints were manipulated: 1) structural-semantic bias, via a verb or adjective-biasing adverb (*very/immediately*), and 2) the frequency (BNC) by which a participle appears in its verbal/adjectival form. A mixed model analysis revealed a main effect of structural-semantic bias for which the direction of the effect differed for native speakers and L2 learners. Native speakers had marginally longer RTs at the first verb region of the rational clause and significantly longer RTs at the verb+ region (*any spies*) for the adjectival passive vs. the verbal passive and a by-phrase control (3). For the L2 learners, both the adjectival and verbal passives had longer RTs than the control, beginning at the verb and through the final regions, suggesting that both forms were processed as adjectivals. No effect of frequency was found for either group.

In sum, these data support the claim that there exists a structural-semantic distinction between verbal and adjectival passives and that for native speakers, only the former can introduce an implicit agent argument to the discourse. In addition, the L2 data suggest that the adjectival form in the L1 may constrain selection to an adjectival form in the L2, which is consistent with the claim that lexical-semantic properties of an L1 may transfer and affect L2 processing, although an additional study is needed to rule out general L2 effects.

Examples
(1) *The important document was very classified to keep any spies from discovering the plan.*
(2) *The important document was immediately classified to keep any spies…*
(3) *The important document was immediately classified by the general to keep any spies…*

References

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology, 50*, 119–148.
Kratzer, A. (1996). Severing the external argument from the verb. In Roryck, J. and Zaring, eds. *Phrase Structure and the Lexicon*. Dordrecht: Kluver. 109-138.
Mauner, G. & Koenig, J. P. (1999). Lexical Encoding of Event Participant Information. *Brain and Language 68*, 178– 184.

The interplay of discourse and structural constraints on referential processing: An ERP study
Nayoung Kwon (Konkuk University) & Patrick Sturt (University of Edinburgh)
nayoung.kw@gmail.com

Previous studies have shown that anaphoric noun phrase resolution interacts with discourse comprehension in on-line sentence processing (Van Berkum et al., 1999, 2003, 2007; Nieuwland & Van Berkum, 2008). For example, referentially ambiguous noun phrases (e.g., *the girl* in a two-girl context) elicited a sustained frontal negative shift (Nref effect) compared to non-ambiguous counterparts. In this study, we investigated whether discourse constraints on anaphor resolution, such as those that result in the Nref effect, interact with syntactic constraints.

We manipulated the number of discourse referents in context sentences (one- vs. two-referent(s)) to investigate a referential ambiguity effect (1). The type of anaphoric expressions (reflexive vs. pronoun) was also manipulated (2). Pronouns and reflexives have different syntactic constraints on distribution; while reflexives should be structurally within the scope of a local coindexed category (i.e., Condition A), pronouns cannot be (i.e., Condition B). Additionally, target sentences always had an adjunct control construction, and the main clause subject was a referent introduced in the preceding context, which thus served as the controller of PRO of the adjunct clause (2). This means that while reflexives should be co-referential both with PRO and the main clause subject, pronouns should be co-referential with neither. This predicts that i) the Reflexive conditions should prefer One-referent context, and the Two-referent reflexive condition will elicit a Nref effect due to referential ambiguity; and ii) the Pronoun conditions should prefer Two-referent context in which PRO and the pronoun can be assigned a different referent. In addition, if the referent of the One-referent condition (i.e., *Tom* in 1) was initially incorrectly assigned to the pronoun, the One-referent pronoun condition will elicit a P600 at the main clause subject position, due to a violation of Principle B (2). To test these hypotheses, we conducted an ERP experiment (22 English monolinguals, 500ms SOA, 200ms ISI, 64 channels).

The results are overall consistent with the predictions. In a latency window of 200-1000 ms post-stimulus onset of reflexives, the Two-referent reflexive condition showed a Nref-like anterior negativity relative to the One-referent reflexive condition ($p < .02$). There was no such effect elicited to the Two-referent pronoun condition. Instead, in a latency window of 300-600 ms post pronoun onset, the One-referent pronoun condition elicited an N400-like negativity in the centro-parietal region compared to the Two-referent pronoun condition. This effect, however, did not reach significance ($p < .096$). In addition, in a latency window of 250-600ms following the main clause subject onset, the One-referent pronoun condition elicited an early anterior negativity in lateral sites ($p < .061$) and a late positivity in centro-parietal regions (midline analysis: $p < .017$) compared to its counterpart reflexives.

Potentially referentially ambiguous contexts (i.e., the Two-referent conditions) elicited different brain responses depending on the anaphor type; the Nref effect was visible only in the Two-referent reflexive but not in the pronoun conditions. This suggests that the referential processing is not only sensitive to discourse context but also to structural constraints. This is also supported by the (very marginal) N400-like effect to the One-referent pronoun condition at critical pronoun; the referential processing of a pronoun could be difficult with a one-referent context, especially when the pronoun occurs in a position that is susceptible to Condition B. Finally, the early anterior negativity and the late positivity to the One-referent pronoun condition at the main clause subject suggest that the referential information is immediately used to parse the rest of sentences (Van Berkum et al. 1999). Thus, once a pronoun is assigned its referent, PRO and the main clause subject are parsed as having a different co-referential index from that of pronouns. Thus, when this prediction is not born out, a processing difficulty results.

(1) It has been a difficult day for Tom/Tom and Jerry.
(2) Last Sunday, PRO_i after calming $him_{*i/k}$/$himself_i$ down doing yoga, Tom_i sang Jerry a lullaby.

Assessing the on-line application of binding constraints without gender stereotype
Kellan Head (Teach for America) & Jeffrey T. Runner (University of Rochester)
jeffrey.runner@rochester.edu

Background. Sturt (2003) used gender stereotyped nouns like *surgeon* to examine the on-line application of binding condition A. First fixations to the reflexive in a sentence like *The surgeon who treated Jennifer/Jonathan had pricked himself/herself* were reliably longer if the reflexive mismatched the stereotyped gender of the accessible antecedent (*surgeon*). Subsequent studies have employed this manipulation to investigate the on-line application of binding condition A in a variety of constructions (Felser et al., 2009; Xiang et al., 2009; Cunnings & Felser, 2012; Dillon, 2012). Though this manipulation has produced intriguing results, it relies on the lesser-understood lexical representation and processing of gender stereotype. In addition, these studies have focused on binding condition A (reflexives), and have generally assumed they will generalize to binding condition B (pronouns).

Experiment. In a visual world eye-tracking study participants (n=17) listened to sentences containing (normed) non-gender stereotyped occupations like *pharmacist* while looking at a grid containing pictures of the mentioned characters and two distractor images. Gender was manipulated visually, using a picture of a male or a female pharmacist. Half of the materials tested reflexives and the other half tested pronouns. See (1)-(2). Participants clicked on the picture corresponding to the anaphor.

Results. The figure below depicts "target advantage" fixations, which are (binding theory accessible) target fixations minus inaccessible fixations. The target advantage was greater for reflexives than for pronouns (p<.001): participants were quicker to fixate the target on reflexive trials. However, starting at about 400ms after the anaphor onset, there is a significant effect of inaccessible gender (p<.001): target advantage is reduced when the inaccessible matches the gender of the anaphor. Listeners are more likely to consider the inaccessible antecedent when it matches the gender of the anaphor than when it does not. The click data were highly accurate (~95% correct) and did not differ by condition.

Discussion. This experiment replicates the basic Sturt finding: condition A seems to influence processing during the earliest moments of reflexive interpretation (<400ms), but adds that condition B is significantly less robust (negative target advantage). Both anaphor types were susceptible to the influence of gender-matching inaccessible antecedents starting at about 400ms post-anaphor. The study contributes to our developing understanding of the on-line application of the binding constraints by using a visual gender manipulation rather than relying on gender stereotype manipulation.

Figure 1.Target advantage fixations on pronoun and reflexive trials with inaccessible gender mis/match.

reflexive gender mismatch   reflexive gender match
ms from anaphor (pronoun/reflexive) onset
pronoun gender mismatch   pronoun gender match
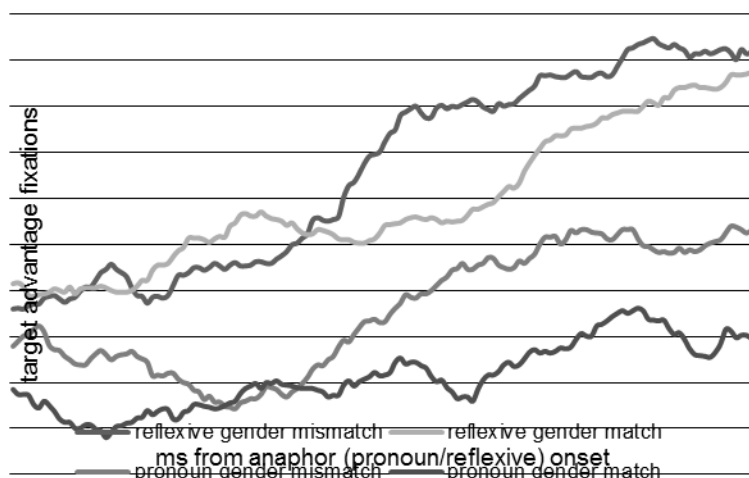
Sample materials

1. *Inaccessible gender mismatch reflexive* (a), *pronoun* (b):
   a. The pharmacist(m) that Molly met drove himself to the party.
   b. The pharmacist(m) that Molly met drove her to the party.
2. *Inaccessible gender match reflexive* (a), *pronoun* (b):
   a. The pharmacist(f) that Molly met drove herself to the party.
   b. The pharmacist(f) that Molly met drove her to the party.

Contextual referent predictability affects optional subject omission in Russian

Ekaterina Kravtchenko (University of California, Santa Cruz)
ekravtch@ucsc.edu

Clause subjects are frequently omitted in spoken Russian [5], in most instances unaccompanied by change in meaning. Omission appears to be mediated by the whether the subject is recoverable in discourse, but is typically optional [2]. A question remains of which pressures on language production this choice is subject to:

1. Maša ostalas' doma, potomu čto ona/Ø zabolela
    Masha stayed.FEM at-home, because she/Ø became-sick.FEM
    "Masha stayed at home, because [she] was sick."

Theories of efficient communication, including Uniform Information Density (UID) [3], propose that speakers optimally trade off efficiency and robustness of the intended message. To this end, speakers would be expected to, where possible, reduce or omit redundant (more contextually predictable) elements. UID predicts that this process should be evident at all levels of production, including in cases of argument omission, pro-drop, and ellipsis [3]. However, it has only been clearly demonstrated for omission of functional elements, e.g. optional *that*-omission [3]. Investigations have also primarily been restricted to English. Previous work suggests that speakers use shorter and less informative referring expressions in English, when referents are more predictable in context [4]. The optionality and frequency of subject omission in Russian provides a test case of whether *omission* of non-functional elements is sensitive to these pressures, as well.

Experiment. 70 native Russian speakers were recruited online to participate in a task where they were asked to guess the identity of upcoming concealed subjects, given only preceding context (method adopted from [4]). Each participant saw 2 of 24 passages from blogs (8), plays (8), or interviews (8), with text progressively revealed as participants indicated whether the upcoming subject was one of any previously mentioned referents, or 'something new.' Only subsequent-mention trials with singular referents were considered (754 subjects; 391 overt/354 null). The contextual predictability of a referent was estimated by computing the negative log probability of participants accurately guessing the concealed subject's identity, in context.

Results. A generalized mixed logit model analysis showed that speakers preferentially omit subjects when they are more contextually predictable ($p < .05$). This effect holds after controlling for coreference with the preceding subject, number of previous mentions, form and grammatical function of last mention, distance to last mention in words, and number of preceding referents; passage was included as a random effect.

Conclusion. This study demonstrates that speakers are more likely to omit subjects when the referent is more predictable given preceding context. The results are compatible with theories of efficient language production, such as UID [3], as well as expectation-driven accounts of reference processing [1]. This suggests that like optional functional elements, omission of non-functional elements, including verb arguments, is influenced by pressure for communicative efficiency.

References
1. Arnold, J. (2008). Reference production: Production-internal and addressee-oriented processes. *Language and Cognitive Processes*, 23(4), 495-527.
2. Franks, S. (1995). *Parameters of Slavic morphosyntax*. Oxford: Oxford University Press.
3. Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, 61(1), 23-62.
4. Tily, H. & Piantadosi, S. (2009). Refer efficiently: Use less informative expressions for more predictable meanings. In *Proceedings of the workshop on the production of referring expressions*, CogSci 2009.
5. Zdorenko, T. (2009). Subject omission in Russian: a study of the Russian National Corpus. *Language and Computers*, 71(1), 119-133.

Poster Session 3

What's in a name? Lexical retrieval during visual object processing
Manizeh Khan, Whitney Fitts & Jesse Snedeker (Harvard University)
khan@wjh.harvard.edu

It is uncontroversial that we sometimes recruit linguistic representations for noncommunicative functions (e.g., the conscious experience of thought as silent speech or verbal rehearsal). How ubiquitous is this phenomenon Prior work suggests that, for infants, simply seeing an object leads to activating its verbal label (Mani & Plunkett, 2010; 2011). On the other hand, studies with adults have had mixed results. Visual-search tasks show effects of homophony between target and distractor images, implicating lexical representations (Meyer, Belke, Telling & Humphreys, 2007). However, there are no effects for onset or rime overlap, speaking against the activation of phonological forms (Telling, 2008). So, the question remains: do adults spontaneously implicitly label visual objects, or does this tendency diminish after infancy? We adapted the infant implicit labeling paradigm (Mani & Plunkett, 2010; 2011) to test whether this simple visual task would also lead to phonological activation in adults.

Experiment 1. 32 Adults were asked to memorize images from 84 three-image trials. Each trial began with the prime image, presented silently for 1500ms, followed by a blank screen for 200ms, and then the Target and Distractor images were presented side-by-side for 2500ms. 50ms after Target-Distractor image onset, the participant heard the label for the Target. On critical trials, the Prime was related to the Target, either via homophony or phonological onset overlap. We compared fixations to the Target on these trials to control trials where the Prime and Target were unrelated. Importantly, for both these conditions, any priming effect would indicate some form of spontaneous label activation for the prime image. A cluster-based permutation test (Maris & Oostenveld, 2007) yielded a significant priming effect in the Homophone condition ($p<.01$), but not in the Phonological-Onset condition ($p>.1$). However, the interaction between the Homophone and Phonological-Onset condition was only marginally significant ($p=.06$), so in Experiment 2 we sought to increase our power to detect any phonological-onset priming effects that might be present.

Experiment 2. To test the sensitivity of this paradigm to phonological-onset priming in adults, we made two changes: 1) doubling the number of phonological-onset overlap trials to 24, and 2) including a between-subjects condition where the prime was auditorily labeled (N=32 per condition). Consistent with Experiment 1, we found no effect of phonological-onset overlap when the prime was not explicitly labeled ($p>.3$). However, when participants heard the prime label, there was a decrease in looks to the target following a phonologically related prime ($p<.01$), and there was a significant interaction between prime-labeling and phonological overlap ($p<.01$). These results indicate that the paradigm was sensitive to phonological-onset priming.

Together, these experiments suggest that simply viewing an image does not lead to phonological activation in adults. This is surprising given that 1) 24-month-olds do show this effect, and 2) there is evidence of lexical activation from the homophone results. One possible explanation is that adults and children both spontaneously recruit verbal labels but the denser lexical network in adults results in weaker phonological priming effects from this implicit activation. Alternatively, during implicit language use, adults may fail to activate phoneme-level representations but still activate representations of lexical form (e.g. lexemes Caramazza, 1997), explaining the discrepancy between homophones and phonological-onset overlap.
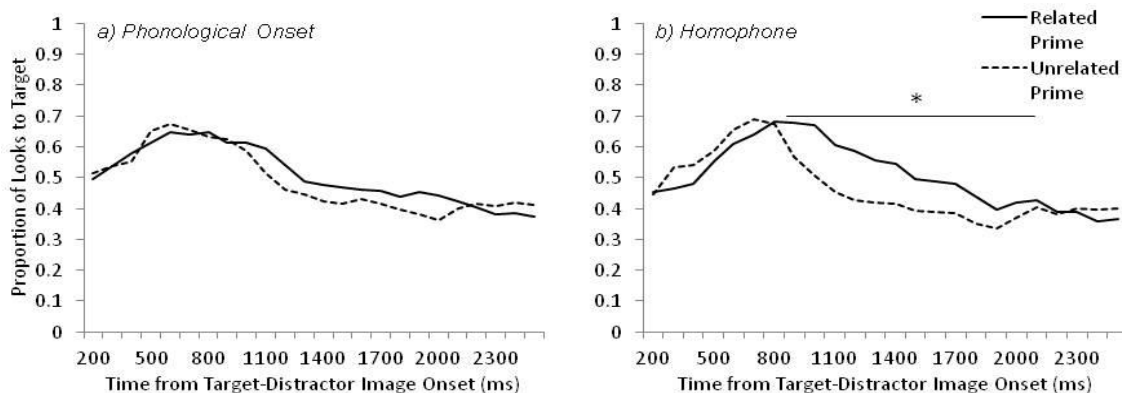
Figure 1. Experiment 1 eye tracking results. Proportion of looks to target image for a) Phonological-Onset, and b) Homophone prime conditions.

Anaphors influence memory for plural antecedents

Nikole D. Patson (Ohio State University)
Huffman.689@osu.edu

There are two ways in which referents of plural noun phrases (NPs) can be conceptually represented: as singular, non-differentiated groups (collective representation) or as multiple distinct entities (distributed representation). Several researchers have shown that plural definite descriptions (e.g., *the cats*) have collective representations while conjoined noun phrases (e.g, *the cat and the dog*) have distributed representations (e.g., Patson & Ferreira, 2009).

Kaup et al. (2002) showed that the choice of anaphor can influence comprehenders' conceptual representations of plural antecedents. In their study, participants read sentences like (1) *John and Mary went shopping*, followed by either (a) *They bought a gift.* or (b) *Both bought a gift.* Afterwards, participants were asked: *How many gifts were purchased?* In the condition with a distributing quantifier (1b) participants indicated that two gifts were purchased. In the condition with a non-distributing quantifier (1a), participants indicated that only one gift was purchased. Kaup et al. argued that this indicates that the pronoun *they* signifies that a collective representation should be adopted, while *both* signifies that a distributed representation should be created.

However, Patson and Ferreira (2009), using a garden-path sentence paradigm, found that the pronoun *they* was capable of picking up the distributed representation of a conjoined noun phrase. They argued that the pronoun *they* does not force a collective interpretation of a conjoined noun phrase, but rather reflects the semantic representation of the antecedent it refers to.

In the current study, we investigated how anaphors influence the memory for plural entities. If an anaphor changes how comprehenders conceptually represent the antecedent, we should find better memory for the antecedent when the anaphor encourages the same type of representation. For example, if the antecedent is a conjoined NP, comprehenders should have a stronger memory representation for the conjoined NP if the anaphor is distributed (both) compared to when the anaphor encourages a collective representation (they).

Participants read short 4 sentence paragraphs. The second sentence of the paragraph contained either a conjoined NP or a plural definite description (as in 1). The third sentence contained either a distributed or a collective anaphor that referred to the plural NP in the second sentence (as in 2). The first and final sentences contained no reference to the plural NP. Participants read the paragraphs and answered comprehension questions. Participants completed an unrelated sentence reading task, and were then given a memory test for the sentences that contained the critical NPs (see 1). In the memory task, participants were presented with either the version of the sentence with the conjoined NP or the version with the plural definite description. Their task was to decide whether the probe was "old" or "new".

1. {Brenda and Todd}/ {The med students} were celebrating at the bar.
2. They/Both called a taxi.

A measure of sensitivity (d-prime) was calculated for each of the four sentence conditions (e.g., Conjoined Antecedent/Distributed Anaphor) to measure participants' sensitivity. There was an interaction between antecedent type and anaphor type. Participants' memory for sentences with plural definite descriptions was not affected by anaphor type. However, for sentences with conjoined noun phrases, participants' memory was more accurate when the anaphor was distributed (*both*) compared to when the anaphor was collective (*they*). The current findings are consistent with Kaup et al.'s findings that *they* encourages a collective interpretation of conjoined NPs. Taken together, both studies indicate that anaphors can influence the conceptual representation of their antecedents.

Competitors chosen by null pronouns in Brazilian Portuguese: Evidence from eye movements
Elisangela Nogueira Teixeira, Maria Elias Soares & Maria-Cristina Fonseca (Universidade Federal do Ceará)
mcrisfon@uol.com.br

Introduction. The choice between a null and an overt pronoun in pro-drop languages has been a subject of intensive research in recent years (Carminati, 2002; Alonso-Ovale, 2002; Fonseca, 2012). Carminati's (2002) findings supported the hypothesis that languages with a pronominal system, with two pronouns, are likely to have different, specialized functions for them, reflecting their use in discourse. Almor (2007) states that the multiplicity of referential forms in languages is not a problem but a solution to overcome the constraints of our memory system. Carminati proposed the Position of Antecedent Hypothesis (PAH), a parsing strategy for intrasentential anaphora, which suggests that the Italian processor assigns more subject antecedent for null pronouns (NP) and more object for overt pronoun (OP). This division of labor has also been reported in intersentential studies in Spanish (Alonso-Ovale, 2002) and in European Portuguese (Costa et al., 1998). Brazilian Portuguese (BP) has been claimed to be a partially null-subject language (Roberts & Holmberg, 2009) due to its recent loss of verbal morphology as well as the sharp increase in the use of OP for 1st and 2nd persons, but not as much for the 3rd person (Duarte, 1996). Empirical data collected in an offline fashion by Fonseca (2012) support the validity of PAH for BP.

Materials & Methods. The present work further investigates the BP by running an eye tracking study with ambiguous intrasentential anaphoras. In order to test the preference of the informants across the competitors (subject and object positions), complex sentences were composed of subordinate clauses having either an overt or a null pronoun in the subject position. In half of the sentences the main clause preceded the subordinate clause (backward anaphora, BA), while in the other half the subordinate clause preceded the main clause (forward anaphora, FA). We constructed 64 experimental items, 16 per type. To ensure ambiguity, two proper names of the same gender were used in the matrix clause. The screen also presented an accompanying question to control the informants' attention and comprehension, as well as the proper names, from which a response should be clicked. The dependent variables are the answers and the total fixation duration of the co-indexed competitors.

Anaphor *vs.* Pronoun		Sentences
Backward	Overt	*Tiago$_i$* waved at *Fabio$_j$*, when he$_{i,j}$ crossed the patio.
Backward	Null	*Tiago$_i$* waved at *Fabio$_j$*, when PRO$_{i,j}$ crossed the patio.
Forward	Overt	When he$_{i,j}$ crossed the patio, *Tiago$_i$* waved at *Fabio$_j$*.
Forward	Null	When PRO$_{i,j}$ crossed the patio, *Tiago$_i$* waved at *Fabio$_j$*.

Results & Conclusion. The results of the online experiment indicate that BP corroborates PAH. Here we investigate the total fixation durations at the competitors (italicized examples above). Significant effects have been detected for the interaction between **null pronouns** in **backward** anaphora *vs.* **competitors** ($F_{1,15} = 7.62$, $p < .015$) as well as for **overt pronouns** in **forward** anaphora vs. **competitors** ($F_{1,15} = 17.9$, $p < .001$). The participants clearly prefer to link the null pronouns to the subject position, regardless being backward or forward anaphora. However, the overt pronoun was more co-indexed with the object position in the backward anaphora.

References
Almor, A., Nair, V. (2007) *Lang. and Ling. Comp.* 1/1-2, 84-99.
Alonso-Ovale et al. (2002) *Journal of Italian Linguistics*, 14:2, 151-169.
Carminati, M.N. (2002) *The processing of Italian subject pronouns*. PhD Thesis. UMass Amherst.
Costa, A., Faria, I., Matos, G. A. (1998) *Actas do VIII ENAPL*. Lisboa. APL, 173-188.
Duarte, M.E.L. (1996) *Do Pronome Nulo ao Pronome Pleno*. Campinas: Ed. da Unicamp.
Fonseca, M.C.M. (2012) *Revista Linguística*. Rio de Janeiro. 8:2, 179-201.
Roberts, I. & Holmberg, A. (2010) *Parametric Variation: Null subjects in minimalist theory*. Cambridge U Press.

Argument identity impacts predictions faster than argument roles

Wing Yee Chow, Cybelle Smith, Glynis MacMillan and Colin Phillips (University of Maryland)
wychow@umd.edu

Previous research has established that the N400 is reduced for more expected words [1], but little is known about how expectations are computed on-line. While a word's predictability is commonly estimated using an off-line cloze task [2], cloze probabilities may not reflect comprehenders' real-time expectations. In particular, a recent study in Chinese showed that (1) the N400 was completely insensitive to a cloze probability manipulation created by reversing the structural roles of the pre-verbal arguments (e.g., cop_{SUBJ} $thief_{OBJ}$ arrest vs. $thief_{SUBJ}$ cop_{OBJ} arrest), and (2) the N400's sensitivity reemerged only when comprehenders had more time for predictive computation [3]. This time-sensitivity argues against an otherwise tempting account based on independent semantic composition [4,5]. Similar cases of N400 insensitivity have been found in verb-final clauses in other languages such as Dutch and Japanese [4,6]. However, these findings seem to contrast with a classic finding about the 'verb final' order created by filler-gap dependences in English, where N400 shows clear sensitivity to the plausibility/cloze probability of a verb (e.g., John knew which {customer|article} the secretary called…) [7]. To reconcile this apparent discrepancy we tested whether argument identity and argument role might impact predictive computations at different time scales. We found that N400 was sensitive to the cloze probability of a verb when argument identity was manipulated, but not when argument role alone was manipulated. We propose that comprehenders can immediately use argument identity, but not argument role, to compute predictions for an upcoming verb.

In the present study we manipulated the cloze probability of the verb (expected: 26% vs. unexpected: 0%) in object-extracted questions in two different ways: in one case cloze was controlled by the identity of the extracted argument (*a* vs. *b*), in the other case by the structural role of an identical pair of arguments (*c* vs. *d*), resulting in a 2 (argument identity vs. role) X 2 (expected vs. unexpected) design. All arguments were animate. If comprehenders can immediately exploit the identity and the structural role of the arguments to anticipate an upcoming verb, then the N400 should be sensitive to the verb's cloze probability in both cases. Alternatively, if predictive computations cannot immediately exploit one or another type of information, then N400 insensitivity might be observed. Results (n=24) revealed that the N400 was sensitive to the cloze difference created by the argument identity manipulation: unexpected verbs elicited a significantly larger N400 and a posterior late positivity (P600), just as in [7]. However, the N400 was completely insensitive to the cloze difference created by the argument role manipulation, despite comprehenders' ability to detect the anomaly, as reflected in a P600 effect. These results show that comprehenders are sensitive to both the identity and the structural roles of the arguments, but that only the argument identity manipulation impacts the N400. This suggests that, at the point when the verb was presented, comprehenders' predictive computations had been impacted by argument identity but not yet by argument role. Together with previous findings in Mandarin Chinese [3], these results suggest that structural role information does not impact comprehenders' predictions immediately, and they resolve the apparent cross-language discrepancy.

Argument Identity
The secretary confirmed **(a)** which illustrator the author had **hired**…
(b) which readers the author had **hired**…
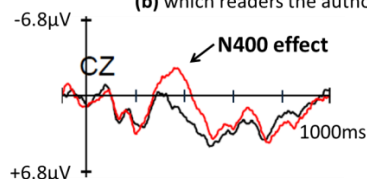
Argument Role
The nanny knew **(c)** which housekeeper the billionaire had **hired**…
(d) which billionaire the housekeeper had **hired**…
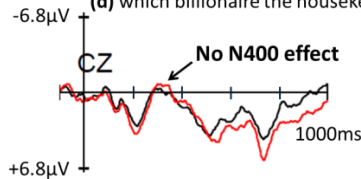
References
[1] Kutas & Hillyard (1984) Nature. 307, 161-163. [2] Bloom & Fischler (1980). Mem. & Cog. 8, 631-642. [3] Chow, Phillips & Wang (2012) CUNY. [4] Van Herten, Kolk & Chwilla (2005) Cog. Brain Res. 22, 241-255. [5] Oishi, H., & Sakamoto (2010) CUNY. [6] Garnsey, Tanenhaus & Chapman (1989) J. of Psycholing. 18, 51-60. [7] Kim & Osterhout (2005) JML. 52, 205–225.

Indefinite NPs introduce new referents but not immediately

Maria Luiza Cunha Lima (Universidade Federal de Minas Gerais), Amit Almor, Evgenia Borschevskaya &
Timothy W. Boiteau (University of South Carolina)
marialuiza.cunhalima@gmail.com

Indefinite NPs such as "a gladiator" are typically interpreted as introducing a new referent into the discourse (Heim, 1988), but the precise time course of this process has not been studied. In particular, it is unclear whether indefinite NPs are interpreted as new references immediately when comprehenders encounter the indefinite article, or whether this interpretation follows an attempt to match the head noun to previously mentioned referents even when preceded by the indefinite article. To better understand the time course of indefinite reference resolution we conducted a visual world experiment comparing indefinite NPs introducing new referents to indefinite NPs co-referring with previously mentioned entities in discourse, as well as a control definite condition. Forty-eight subjects heard two-sentence discourses, which appeared in one of three conditions as illustrated by the sample item below. In the INDEFINITE-NEW condition, the indefinite NP, "a gladiator", introduces a new referent to the discourse. In the INDEFINITE-CO-REFERENTIAL condition the absence of a finite verb indicates that the indefinite NP co-refers with the "gladiator" mentioned in the first sentence. In the DEFINITE-CO-REFERENTIAL control condition, the definite NP "the gladiator" is clearly anaphoric.

Visual displays showed scenes depicting the event described by the discourse with four characters related to the description: an Agent (lion), a Target (gladiator with a sword and a shield), a Competitor (gladiator with a sword and no shield), and an Unrelated character (horse). The scene made clear which character is the Target mentioned in the first sentence (see Figure). Our main interest was in looks at the Competitor in Sentence 2, which we took to reflect consideration of a new referent. We used Dynamic Generalized Linear Models (West and Harrison, 1999) and report here the significant effects.

At the beginning of Sentence 2, subjects did not look at the Competitor in either condition. Differences between the conditions in looks to the Competitor emerged 200 ms after the offset of the verb ("killed") in the INDEFINITE-NEW condition, at which point there were more looks to the Competitor in that condition than in the other two conditions. There were more fixations to the Competitor in the INDEFINITE-NEW condition than the INDEFINITE-CO-REFERENTIAL condition up to 400 ms after the end of the relative clause. There were no differences in looks to the Competitor between the INDEFINITE-CO-REFERENTIAL and the DEFINITE-CO-REFERENTIAL at any time point. In sum, these results indicate that subjects looked for a new referent only in the INDEFINITE-NEW condition, and even in that condition this process started only 200 ms after processing the verb (allowing 200 ms for programming the eye movement).

Overall these results show that the indefinite article does not induce the introduction of a new reference in and of itself. This may indicate that introducing new referents is costly, and is therefore postponed until more information (in this case a verb) unambiguously indicates that a reference to a new referent was made.

The lion attacked a gladiator in the arena. **A gladiator** who had a sword killed the lion.	INDEFINITE-NEW	
The lion attacked a gladiator in the arena. **A gladiator** who had a sword and a shield.	INDEFINITE-CO-REFERENTIAL	
The lion attacked a gladiator in the arena. **The gladiator** who had a sword killed the lion.	DEFINITE-CO-REFERENTIAL	

References
Heim, I. (1988). *The semantics of definite and indefinite noun phrases.* New York: Garland Pub.
West, M., & Harrison, J. (1999). *Bayesian Forecasting and Dynamic Models.* Springer Series in Statistics.

The cost of unexpected contrast: Processing 'let alone'
Jesse A. Harris (Pomona College)
jesse.harris@pomona.edu

In *let alone* constructions, e.g., *John can barely run a MILE, let alone a MARATHON*, contrastive focus marks the elements under comparison, which form propositions standing in a scalar relationship. Thus, the processor faces two potential costs: (i) recovering the appropriate propositions for comparison and (ii) placing those propositions on a contextually relevant scale. This study reports two main findings regarding those costs: first, recovering the proposition from a VP remnant for comparison (2b) is less taxing than from a DP remnant (2a), and, second, the scalar operator *even* facilitates the scalar inference associated with *let alone*. In addition, *even* merely signals a scalar relationship, not focus placement, as would be expected if the processor treated it on par with the better-studied focus operator *only*. These findings are not expected given previous literature on focus sensitive elements. I propose that the results reflect a general preference for broad focus by default, and that focus sensitive particles without truth-conditional consequences like *even* do not overturn that preference.

Previous reading studies on corrective ellipsis constructions lead to two main conclusions. First, in corrective stripping constructions, e.g., *Jane passed the salt to her mother, but not her father*, remnants (*her father*) preferentially contrast with the most deeply NP (*her mother*). Second, this preference can be modulated by elements marking contrast overtly, e.g., by a focus sensitive operator (FSO) like *only* (Paterson et al, 2007; Stolterfoht et al, 2007) or context strongly constraining focus placement (e.g., Carlson, 2011). Crucially, in these studies, FSO *only* is assumed to mark the constituent to its immediate right with contrastive accent (Buring & Hartmann, 2001) by default. Unlike *only*, scalar *even* does not strongly correlate focus prominence with syntactic position. As such, the presence of *even* might be less informative than *only* with respect to marking focus prominence in silent reading, so that the processor remains uncommitted to the exact placement of stress. Thus, the prediction is that the processor will not immediately commit to focus placement after *even*, in contrast to findings observed for *only* (Paterson et al, 2007; Carlson, 2011).

In a sentence completion task, subjects (N=26) completed sentences after *let alone* with or without *even* as in (1). As predicted, sentence completions showed a general preference for VP completions (e.g., *read it*) (65%), z=2.20, p<0.05, with the remainder resulting in DP completions (e.g., *the book*). However, *even* did not boost VP remnant completions (d=2%), as would be expected if FSOs uniformly attract contrastive focus to the immediate right, in this case the verb (*skim*). The preference for VP contrast after *let alone* was then tested in an eye movements study (N=36), crossing remnant type (VP vs. DP) with presence of *even* (2), with items independently normed for naturalness (N=22). Evidence for a VP preference in reading after *let alone* was supported by a variety of measures, including shorter go past times and fewer regressions out for VPs (2b) over DP counterparts (2a) on the remnant (region 4). Evidence for facilitation following *even* was observed in fewer regressions out from the remnant, as well as fewer regressions in and shorter second pass times on the region containing *even* (region 2). Again, *even* did not facilitate VP remnants over DP remnants, as would be predicted if it assigned narrow focus to the neighboring element, e.g., the matrix verb (*skim*).

(1)   Pat didn't (even) skim the article, let alone _____.

(2)   a. $|_1$ Pat didn't $|_2$ (even) skim the article, $|_3$ let alone $|_4$ the book, $|_5$ before his class $|_6$ …
            b. $|_1$ Pat didn't $|_2$ (even) skim the article, $|_3$ let alone $|_4$ read it, $|_5$ before his class $|_6$ …

Effects of novelty and givenness on acoustic reduction
Lap-Ching Keung & Jennifer E. Arnold (University of North Carolina, Chapel Hill)
lkeung@unc.edu

Repeated words within a discourse tend to be acoustically reduced, i.e. shorter in duration and less clearly articulated (Fowler & Housum, 1987). This variation may result from a pragmatic selection rule, where the speaker uses reduced forms for given referents and acoustically prominent forms for new referents (Halliday, 1967). Or it may be due to the speaker's ease of producing a word (Kahn & Arnold, 2012), since previously mentioned words are easier to retrieve and produce. An open question is whether effects from discourse status can be separated from effects of facilitation. We tested this question with a production experiment. Participants viewed an image of four objects and described a sequence of two movements (e.g., *The beetle moved above the ruler. The <u>lemon</u> moved above the carriage).* The target was the object in the second utterance. We analyzed onset latency and target duration.

We manipulated the <u>discourse status</u> of the target: nonrepeated targets were "new"; repeated mention targets were "given". We also manipulated <u>processing difficulty</u> by contrasting familiar words (*beetle*, *ruler*) with novel, recently learned words (*beeler*, *roodle*). Participants learned these words as the names for novel pictures. Novel object names were extremely low frequency and were accompanied by unfamiliar conceptual referents, both of which are likely to increase the difficulty of word retrieval and planning. Participants saw each target twice, one condition in the first half of the experiment, the other in the second.

Sample stimuli

Our experiment used a novel paradigm that de-confounded the tendency for low-frequency words to be hard to retrieve and produce from the fact that low frequency words tend to have a longer inherent word length (Zipf, 1935). The novel words were all recombinations of syllables from the familiar words, thus matching their average phonological lengths. That, along with matched biphone frequency and syllable structure, ensures that any observed differences in word duration are the result of familiarity and ease of retrieval. We predicted longer durations for new vs. given, for novel vs. familiar words, and for first vs. second half. The critical question was whether givenness effects were affected by facilitation: 1) Did givenness and novelty interact? 2) Was either effect mediated by processing facilitation, as indexed by latency and speech rate?

Preliminary Results (data from 12 of 24 participants). Log-transformed onset latencies and durations were analyzed. All F1's(1,11) and F2's(1,30) > 4, p's<0.05, unless indicated. Raw durations are shown for interpretability. For both onset latency and target durations, there were main effects of novelty and givenness, such that durations were longer for novel vs. familiar, and for new vs. given. There was also a marginal interaction (*onset*: p1=0.11; p2=0.09; *object*: p1=0.09, p2=0.60). Latencies and novel target durations were also shorter in the second half. Our results suggest that givenness effects on acoustic reduction are partially mediated by processing facilitation: 1) givenness had a greater effect on novel targets, which were harder to retrieve and produce; 2) onset latencies patterned with target durations. However, givenness was not entirely a facilitation effect. In a follow-up analysis, we normalized target durations as a proportion of total utterance time, to assess relative prominence. In this analysis, duration was primarily predicted by givenness, and not novelty. This suggests that the use of reduced forms for given targets persist over and above the concomitant effect of givenness on facilitation.
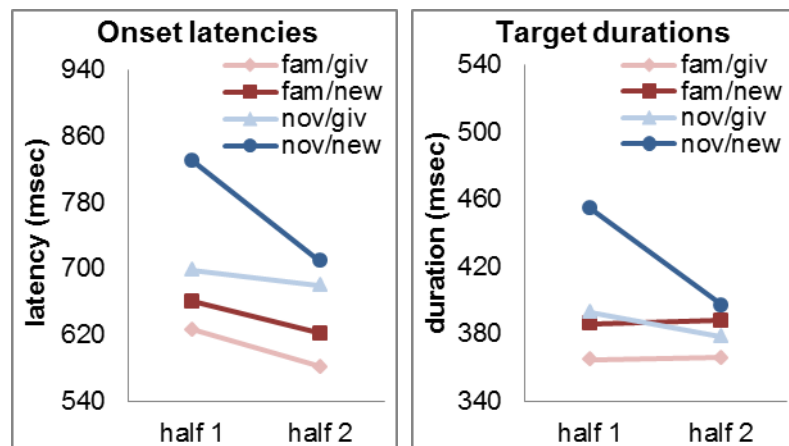
Auditory priming affects planning and execution separately
Jason Kahn & Jennifer Arnold (UNC – Chapel Hill)
jmkahn@email.unc.edu

Acoustic variation in word pronunciation occurs at least in part because of the facilitation of production processes, either from repeated mention of a word (Bard et al., 2000, *JML*) or from contextual predictability (Bell et al., 2009, *JML*). Recent work also shows that pre-planned words sometimes exhibit shorter durations, suggesting a link between planning facilitation and duration (Gillespie, 2012, *Ph.D. thesis*). Duration does not always pattern with latency, however (Kahn & Arnold, 2012, *JML*; Schriefers et al., 1990, *JML*), raising questions about the conditions under which planning does or does not explain facilitation effects on duration. Here we investigate these questions by examining two effects: 1) facilitation from identity primes, and 2) inhibition from semantically-related primes. If their effects on duration and latency pattern together, it would suggest that these facilitatory and inhibitory effects are the result of pre-planning.

We test this question with a variant of the PWI paradigm, in which participants describe simple object movements, e.g. "the airplane rotates". We presented primes at different SOAs, relative to the target stimulus, to examine the time course of their effects. We consider utterance onset latency to be a rough measure of the utterance's planning time, up to the point necessary for fluent initiation. Semantically-related primes tend to elicit longer naming latencies in PWI studies (e.g. Schriefers et al., 1990), arguably due to the difficulty of processing a related lexical item during word production. Identical primes, by contrast, elicit shorter naming latencies, presumably by facilitating the reactivation or reselection of the same information (Glaser & Dungelhoff, 1979, *JEP:HPP*). Semantic effects tend to appear primarily at SOAs shortly before the target stimulus, while identity effects appear in a wider range, both before and after the target (Glaser & Glaser, 1989 *JEP:G*). Participants described the movement of simple colored objects. Prior to each object's movement, participants heard one of four primes: a baseline condition of silence (**C**ontrol condition), the name of the object (**I**dentical condition), predicted to have facilitatory effects, or a semantically-related (**R**) or -unrelated (**U**) word, both predicted to have inhibitory effects. Primes could appear at one of five SOAs: 300, 600, 900, 1200, or 1500ms prior to the object appearing. All participants saw 100 items in randomly-chosen condition and SOA combinations, and across participants each item appeared an equal number of times in each cell. Duration and latency were analyzed with a multi-level model with appropriate random intercepts, slopes, and control variables (including a proxy for speech rate), and Tukey multiple comparisons examined the effect of condition overall and within each SOA.

The primary analysis examined all data, pooling across SOAs. Duration Results: **I** elicited shorter durations than all other conditions, including **C**. This replicates the standard repeated word effect. **R** and **U** did not differ, both eliciting longer durations than **C**, suggesting that pronunciation was slowed by interference from recently processing of any word-like stimulus, not only a semantically-related word. Latency results: **I** and **C** conditions did not differ, demonstrating that the facilitation from the identical prime did not facilitate planning. Both **R** and **U** were longer than the other conditions, but not different from each other. This suggests that non-identical lexical items interfere with planning, mirroring their effect on duration. A subsequent analysis examined condition effects broken down by SOA. Despite its lower power, it revealed the same latency effects at all SOAs, and similar duration effects at SOA300. Semantic inhibition (**U** < **R**) also appeared at SOA300 and 600, but only for latency.Our results demonstrate that duration variation is influenced by both a) identity facilitation, and b) lexical inhibition. Although lexical inhibition affected pre-utterance planning and duration similarly, the effect of identity priming was not linked to planning facilitation. This shows that the mechanism for identity priming does not reduce solely to planning.

PRIME	(airplane)		(silence)		(cinch)		(cloud)
TARGET DURATION	**I**dentical	<	**C**ontrol	<	**U**nrelated	=	**R**elated
LATENCY	**I**dentical	=	**C**ontrol	<	**U**nrelated	=	**R**elated

Table 1. Pattern of significant condition differences in the primary analysis.

Form interference effects during silent reading
Iya Khelm, Naoko Witzel & Jeffrey Witzel (University of Texas at Arlington)
iya.khelm@mavs.uta.edu

This study investigates the role of phonology in silent reading. Although studies have shown that phonological repetition causes reading comprehension difficulties (e.g., Acheson & MacDonald, 2011), the nature of these processing costs remains unclear. The present study attempted to shed light on this issue by examining (i) the time-course of form-related interference effects, (ii) the interaction of these effects with syntactic processing difficulty, and (iii) the extent to which these effects relate to phonological or orthographic overlap.

Participants' eye movements were recorded as they read reduced and unreduced relative clause (RC) sentences like the following:

> (1a) **unreduced/reduced RC; O+P+**
> The infection (that was) left **by the <u>injection</u>** badly hurt the young child.
> (1b) **unreduced/reduced RC; control**
> The infection (that was) left **by the <u>medicines</u>** badly hurt the young child.
> (2a) **unreduced/reduced RC; O+P-**
> The laughter (that was) caused **by the <u>daughter</u>** continued on throughout the party.
> (2b) **unreduced/reduced RC; control**
> The laughter (that was) caused **by the <u>minister</u>** continued on throughout the party.

Experiment 1 (EX1; *N*=32) tested sentences as in examples (1a) and (1b); Experiment 2 (EX2; *N*=32) tested sentences as in (2a) and (2b). As illustrated in these examples, the *by*-phrase in the RC contained (i) a word that overlapped in form with the head of the subject NP (*infection… **injection***) or (ii) a length and frequency matched control word (*infection… **medicines***). Form-related words differed from the subject head by only one letter. In EX1, these words were also strongly phonologically related to the head noun (O+P+), differing from it by only one phoneme (*infection… **injection***). In EX2, these words were phonologically dissimilar to the head noun (O+P-; *laughter… **daughter***). It was predicted that if phonological overlap influences early stages of comprehension, form-related processing difficulty should be revealed in first-pass reading time (RT) measures and should occur independently of syntactic processing effects. However, if phonological overlap primarily affects retention and retrieval in working memory, form-related processing difficulty should be obtained only in later RT measures and should interact with syntactic processing difficulty. Finally, form overlap effects that are truly phonological in nature should be obtained in O+P+ sentences, but not their O+P- counterparts.

Consistent with previous studies (e.g., Ferreira & Clifton, 1986), both experiments yielded robust indications of syntactic processing difficulty at the disambiguating *by*-phrase in reduced RC sentences. Both experiments also revealed processing difficulty at the form-related word – the O+P+ word in EX1 and the O+P- word in EX2 – under first-pass and total RT measures. These form interference effects did not interact with syntactic processing difficulty in first-pass RT measures. In EX1, however, there was an interaction of form overlap and RC type under total RT, indicating particular processing difficulty for reduced RC sentences with O+P+ words.

Taken together, these findings indicate that form similarity causes interference even in the early stages of silent reading. These effects cannot be attributed exclusively to phonological form overlap, as processing difficulty was found for both O+P+ and O+P- words. However, the late interaction of form overlap and syntactic processing effects only for O+P+ sentences suggests that phonology, but not orthography, plays a key role in comprehension processes related to retention and retrieval in working memory.

It's probably porridge: The role of tonal probability in Mandarin lexical access

Seth Wiener & Kiwako Ito (The Ohio State University)
wiener.24@osu.edu

Word recognition in Mandarin Chinese involves processing of both segmental and suprasegmental information. Previous research has suggested that segmental information holds a perceptual advantage over tone [1] and that tone behaves like lexical stress both in production and perception [2]. More recent work has concluded that tonal and segmental information are accessed concurrently and play comparable roles in lexical activation [3]. All previous work, however, assumes the contribution of tone is static across all syllable types, without taking into account the vast disparities of both tonal probability and segmental frequency. The current study investigates whether syllable-specific tonal probabilities guide online lexical access in native Mandarin listeners.

The experiment manipulated syllable frequency (high/low) and tonal probability (most probable/least probable) based on [4]. Participants' eye-movements were monitored while they searched for Chinese characters that matched spoken words. Each critical slide showed four characters: the target (zhōu 'porridge'), a cohort competitor with the identical segment but with a tone that was either more or less probable than the target (zhóu 'axle'), a rhyme competitor with the same vowel-tone combination as the target (gōu 'hook'), and an unrelated distractor with a comparable syllable frequency (juàn 'silk fabric'). Filler trials presented either semantically related or visually similar character pairs. Number of strokes (i.e. visual complexity) and individual character frequency were controlled. If Mandarin speakers are sensitive to syllable frequency and also track tonal probabilities across different syllables, they may more accurately and swiftly detect frequent syllables with highly probable tones than infrequent syllables with least probable tones. In contrast, if syllable-specific tonal probabilities are not part of the phonological representation of the Mandarin lexicon or their contribution to word recognition is minimal, the degree of competition between the target and the cohort competitor should be equivalent regardless of the target's tonal probability.

Responses from 48 participants revealed the main effect of tonal probability and the interaction between tonal probability and syllable frequency on both the fixation likelihood and mouse response times. The main effect of syllable frequency was not significant. A post-hoc comparison revealed that low frequent syllables with most probable tones were fixated and responded fastest while low frequent syllables with least probable tones were fixated and responded slowest. Further analysis showed a significant effect of tonal probability on both the fixation likelihood and the response times only for the low frequent syllables but not for the high frequent syllables. Due to the inherent tonal homophony present in highly frequent syllables, tonal probability may work in tandem with discourse context to further reduce the number of lexical candidates. These results indicate that syllable-specific tonal probabilities are stored, at least for low frequent syllables, and that they affect the processing of words containing these infrequent syllables.

References
[1] Cutler, A. & Chen H-C. (1997). Lexical tone in Cantonese spoken-word processing. Perception & Psychophysics, 59 (2) 165-179.
[2] Chen, J-Y., Chen, T-M., & Dell, G. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. Journal of Memory and Language, 46, 751-781.
[3] Malins, J. G., & Joanisse, M. F. (2010). The roles of tonal and segmental information in Mandarin spoken word recognition: An eyetracking study. Journal of Memory and Language, 62, 407-420.
[4] Cai, Q. & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. Plos ONE, 5(6), e10729.

Effects of context and individual differences on processing taboo words within sentences

Adina Raizen, Cassie Palmer-Landry, & Kiel Christianson (University of Illinois at Urbana-Champaign)
raizen1@illinois.edu

Many studies employing a variety of experimental methodologies have found that taboo words are processed differently than neutral words. Until recently, the processing of taboo words has been investigated primarily using unnatural paradigms that involve the rapid presentation of individual words. A few studies have begun to use more naturalistic, sentence-reading paradigms but still use offline memory measures only. The present study extends previous research by using online eye movement measures to investigate the differences in processing taboo (e.g., *shit*) versus non-taboo words (e.g., *rats*), as well as offline accuracy and reaction time measures on word probes. Another novel aspect of this study is the use of individual differences in offendedness to determine readers' attitudes toward people who swear and how offensive readers found taboo words; we predicted that individual differences would affect how the taboo words were processed.

Previous studies using rapid presentations of individual words have led to two main theories regarding the processing of taboo words (outlined by MacKay et al., 2004). The first is global resource theory, which states that taboo words sap attentional resources and garner more attention than surrounding words. The second is binding theory, which states that the presence of taboo words increases memory for certain aspects of the context specific to the taboo word, but not surrounding words. The present study aims to determine if and how these two theories interact.

A fully crossed, factorial experiment manipulated three variables, resulting in a 2x2x2 Latin square design: inappropriate vs. appropriate situations in which to use a taboo word, expected vs. unexpected users of taboo words ("sinners" and "saints"), and taboo words vs. non-taboo words (see example below). Taboo and non-taboo words were controlled for frequency and length. Speakers and situations were normed for relative offensiveness. Eye movements were recorded while participants (N=80) read sentences silently. After each sentence, a probe appeared asking if they had seen a word in the previous sentence. The probe word was always the word preceding the (non)taboo word (e.g., "shouted," "said"). Participants also completed an offendedness survey.

Linear mixed effects modeling demonstrated that online measures support global resource theory: overall, fixations were longer on taboo words than non-taboo words, and longer fixations on taboo words were predicted by situation and speaker such that when "saints" were in situations inappropriate for swearing, participants spent more time processing taboo words. Offline probe word recall, however, supports binding theory: accuracy increased and reaction time decreased on trials involving a taboo word. Individual differences of offendedness show that participants who were more offended by taboo words looked longer at the probe word in the sentence. They also had *lower* accuracy when recalling the probe word, implying that the fixation on the probe word was used as more of an avoidance technique in order to delay reading the upcoming taboo word, and that the word was not being fully processed. Results imply that global resource theory is accurate to an extent. Taboo words command more attention than neutral words, but this attention actually facilitates binding mechanisms for the surrounding context as predicted under binding theory. However, this general pattern is modulated by individual differences in the degree to which people find taboo words to be offensive, and can actually lead to a disruption of binding mechanisms. Together these findings demonstrate the importance of investigating word processing and memory within a naturalistic sentence-reading paradigm.

Example Sentence

	Speaker (Saint/ Sinner)	Situation (Appropriate/ Inappropriate)	Probe Word	Exclamation (Taboo/ Nontaboo)	
The	chauffeur/cabbie	was going home and/ was driving the old lady and	shouted,	"Douche!"/ "Doofus!"	when he got cut off.

Topic, empathy, and point of view

Laura Kertz and Corey Cusimano (Brown University)
laura_kertz@brown.edu

Topics play an essential role in structuring language at both the sentence and the discourse level. A topic is generally understood to be the entity that a sentence or discourse is about, a characterization that links the function of topical expressions (as devices which support continuity in discourse) with their representation (as both cognitively salient and linguistically prominent). An aspect of topic-hood which remains less explored, however, is the role that topics play in structuring point of view (but see, e.g., Gernsbacher et al. 1992 for related discussion of protagonist effects in story comprehension).

Kuno's (1987) 'empathy' model characterizes a topic as the entity whose perspective and mental states are most aligned with the speaker's for the purposes of structuring an utterance. Simply put, any event can be described from multiple points of view; establishing a topic means favoring one participant's point of view above others'. In support of this, Kuno contrasted pairs like (1) and (2), which report events that are truth conditionally equivalent, but differ in acceptability. Kuno argued that the problem with (2) is that, in the first clause Mary is established as the topic but, in the next, Mary is explicitly marked as a non-topic (appearing in an oblique 'by' phrase), generating a conflict across the two clauses.

(1) **Mary** had quite a time at the party last night. At the end of the night, **she** propositioned the bartender.

(2) # **Mary** had quite a time at the party last night. At the end of the night, the bartender was propositioned **by her**.

(3) **Mark** had not spoken with the new landlady until
   a. **he** contacted her about the deposit.
   b. she contacted **him** about the deposit.
   c. **he** was contacted by her about the deposit.
   d. she was contacted **by him** about the deposit.

As a test of the empathy hypothesis, we constructed 16 stimulus sets as in (3), manipulating both the order of mention for referents in the second clause and whether the second clause was expressed as an active or passive. The crucial comparison involves (b) and (d). In both cases, a 3sg masculine pronoun (him) appears in a non-subject position and co-refers with the subject/topic ('Mark') of the previous clause. Crucially in (3d) the pronoun appears in an oblique by-phrase, explicitly marking it as a non-topic. Under the empathy hypothesis, re-evoking the topic 'Mark' from this position should result in a violation.

Materials like (3) were tested in separate acceptability (n=45) and eye-tracking while reading (n= 48) studies. Acceptability ratings showed an interaction (p<.05) where the (d) sentences were dispreferred compared to other conditions. Reading analyses showed effects of the passive manipulation with increased first-pass (p<.05), regression (p<.001), and total dwell (p<.001) times at object pronouns ('him'/'her'). The predicted interaction between voice and order of mention was observed at the preceding verb ('contacted') with elevated dwell times (p<.01) for (d) conditions.

These results suggest an overall cost associated with passives, consistent with prominence-based models of nominal co-reference, while the observed interaction, which is not predicted under such accounts, confirms that a conflict in point of view can both reduce acceptability and disrupt processing. Moreover, the fact that the interaction emerged only in late measures indicates that the topic manipulation did not increase the difficulty of resolving pronouns (which were unambiguous), but instead disrupted integration across clauses.

References

Gernsbacher, M. A., Goldsmith, H. H., & Robertson, R. R. W. (1992). Do readers mentally represent characters' emotional states? Cognition & Emotion, 6(2), 89–111.
Kuno, S. (1987). Functional syntax: Anaphora, discourse and empathy. University of Chicago Press.

Frequency and distribution of *some* (but not *all*) implicatures
Judith Degen, Michael K. Tanenhaus, Christine Gunlogson (University of Rochester)
jdegen@bcs.rochester.edu

Scalar implicatures can arise from utterances like (1), standardly taken to have a semantic lower-bound interpretation (*at least one student failed*) and a Gricean upper-bound interpretation (*not all of the students failed*). Not all occurrences of *some*, however, so easily lend themselves to the upper-bound interpretation. *Some* can appear as an indefinite determiner with singular head nouns, for instance, where it does not give rise to a scalar implicature. Between these relatively clear-cut cases lie many uses of *some* that are harder to categorize. Is a scalar implicature warranted with mass/abstract nouns as in (2)? In (3), a *not-all* interpretation is plausible, but it does not seem to be the central point.

(1)   Some of the students failed.
(2)   The courts generally show some leniency to first-time offenders.
(3)   Our program will discuss some causes of grief.

The literature on scalar implicatures – both theoretical and experimental – has concentrated on cases like (1), with the (often implicit) assumption that scalar inferences are nearly obligatory when *some* is encountered (e.g., Huang & Snedeker, 2009). However, this Frequency Assumption is untested to date. Its validity has far-reaching consequences. For example, delays in implicature processing have been attributed to a staged process of pragmatic inference: first the semantics of the utterance is computed, and pragmatic enrichment follows in a second costly step. However, these results might instead reflect a frequency effect within a constraint-based system (Degen & Tanenhaus, 2011). Robustness and speed of computing a scalar implicature might be a function of degree of probabilistic support from contextually available cues. We present two studies that combine corpus analyses with rating studies to take a first step towards the methodologically challenging task of identifying and quantifying the frequency of, and the factors that affect, upper-bound interpretations in naturally occurring utterances

Study 1. We extracted all 1951 utterances containing *some* from the Switchboard corpus. 561 cases where the head noun of the *some*-NP was a singular count noun were excluded because these are cases of specific indefinites. We collected 10 implicature ratings for each item on Mechanical Turk. Participants saw the printed target utterance preceded by 10 lines of dialog context. They were then presented with a minimally different utterance where *some* was replaced by *some, but not all*, thus spelling out the otherwise implicit scalar inference (e.g., *Some of the students failed* was paired with *Some but not all of the students failed*). Participants rated the similarity of the two utterances on a 7-point scale: 1 = "very different meaning"; 7 = "same meaning". The mean similarity rating was only 3.7. Moreover, of the 42% of utterances rated either 1 or 7, only 39% were rated 7. Only 41% of ratings were higher than the midpoint 4; 50% were lower than 4. In sum, contrary to the standard assumption, most uses of *some* do not strongly support a scalar implicature..

Study 2. To further explore factors that mediate an implicature, we used the same dataset to extract information about: a) whether *some* occurred in the partitive; b) the NP head's discourse accessibility (discourse-old or new); c) the NP's topicality as measured by whether it was topicalized/in subject position vs. in object/other positions. Cue choice was based on theoretical considerations. A mixed effects linear regression predicted similarity rating from fixed effects of partitive, discourse accessibility, and topicality. Partitive NPs, discourse-old heads, and topical NPs each made the similarity rating significantly higher.

These results suggest that contrary to what is standardly assumed, *some* does not typically provide strong support for an implicature. More generally, the work reported here is an important first step towards understanding how implicature rates are modulated by a variety of linguistic factors - an enterprise that is a necessary prerequisite for designing and interpreting experimental studies of how implicatures are computed in on-line processing.

Effects of speaker identity on processing rude and polite language: Evidence from a Twitterish paradigm

James Nye, Steven Luke, Justine VanDyke-Lyon, Fernanda Ferreira (University of South Carolina)
Jim.m.nye@gmail.com

Traditional theories of language comprehension generally distinguish linguistic sources of information such as syntax and sentence semantics from nonlinguistic sources such as speaker identity. However, according to recent research, mechanisms of language comprehension consider linguistic input and speaker-relevant information simultaneously during online processing. If the content of an utterance is expected given the identity of the speaker, processing is facilitated at both lexical and sentential levels (Van Berkum et al., 2008). The goal of this research was to test this further by examining the processing of profanity produced by different types of individuals. Some people are generally viewed as polite and therefore unlikely to use profanity (e.g. Oprah Winfrey) whereas others are viewed as more provocative and therefore more likely to use profanity (e.g. George Carlin). Our experiments examined whether inferences about the speaker affect the processing of profanity relative to neutral synonyms. Our previous work examined people's reading of sentences containing taboo words (e.g. bitch) versus semantically neutral synonyms (e.g. whine) equated on length. We observed no differences in processing time on any eye movement measure, and no changes in those measures across experimental trials. These findings suggest that when the overall meaning of a sentence is controlled, taboo and neutral words are processed similarly. However, these sentences were shown in typical psycholinguistic fashion - they appeared in isolation, one after another on a computer screen, with no information provided about the person who produced the item. The current experiment used a Twitter-like format to provide information about the "speaker" together with the sentence. The prediction was that neutral and profane words would be processed differently depending on compatibility between the word and speaker.

Participants were shown 72 displays in which a sentence was paired with an "avatar"-like picture meant to be interpreted as the speaker. 24 were experimental items and the rest were fillers (contained slang but no profanity). The design of the experiment was a 2x2: the speaker appeared either polite or rude, and the sentence either contained a neutral or a profane word. Sentence and avatar were shown together during the trial, in a Twitter or Facebook-like manner:

Polite speaker - Neutral content   Rude speaker - Profane content
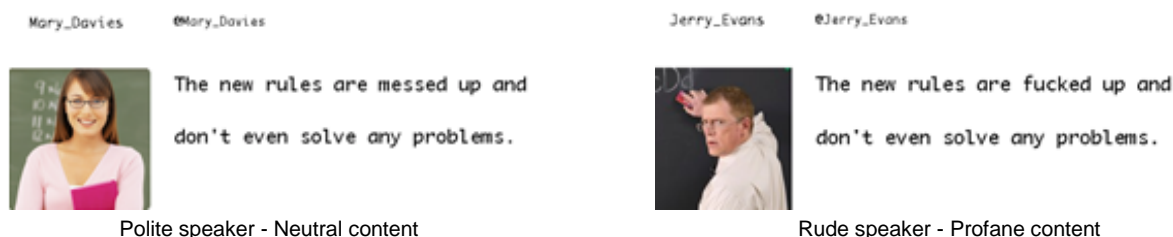
The findings were as follows: First, as expected, fixations times on the speaker image were long in only one of the four conditions - the one that paired the polite speaker with the profane content. Second, reading times on the profane versus neutral word indicated that participants spent more time on likely combinations of word and speaker - they fixated longer on neutral words that were paired with polite speakers, and profane words that were paired with rude speakers. These results suggest that speaker identity has a profound effect on sentence processing, particularly when they are contrasted with results obtained from the more traditional reading experiment that examined profanity. Fixation times on words differ depending on whether those words are likely to be used by the speaker. Taken together, the experiments reinforce the importance of studying pragmatic / sociolinguistic variables such as speaker identity. They also suggest that profane words differt from other lexical items in that they convey little semantic content, and instead are markers of speaker attitude toward the subject of the sentence.

Reference. Van Berkum, J. J., Van Den Brink, D., Tesink, C. M., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of cognitive neuroscience, 20*(4), 580-591.

Facial feedback and the real time comprehension of emotional language

Seana Coulson, Joshua Davis, Piotr Winkielman (UC San Diego)
scoulson@ucsd.edu

Embodied theories of language comprehension suggest language comprehension involves the reactivation of neural systems involved in experience with the domains referred to by a given sentence. This predicts that facial feedback from the production of emotional responses (e.g. frowning) might impact the difficulty of understanding sentences about emotional events. Accordingly, Havas and colleagues covertly manipulated participants' facial expressions as they made valence judgments on sentences about pleasant and unpleasant events (Havas, D. A., Glenberg, A. M., & Rink, M. (2007). *Psychonomic Bulletin and Review* **14**, 436-441). Consistent with the idea that emotional systems contribute to language comprehension, they found that reading times were longer for sentences that were incongruent with the participant's facial expression (e.g. unpleasant events while smiling) than congruent (e.g. pleasant events while smiling). However, results reported by Havas et al. (2007) might reflect the impact of facial feedback on either the comprehension of language, decision processes related to the valence judgments, or both. Further, even if it were clear that facial feedback impacted language comprehension, their use of whole sentence reading times leaves open the precise timing of those effects as well as the linguistic level at which they occur. Here we tested whether facial expression impacts the brain's real time response to emotional language using scalp-recorded event related potentials (ERPs).

Facial expression was manipulated by asking participants to hold chopsticks either in their LIPS ONLY (allowing relatively normal movement of muscles used for smiling) or in their TEETH AND LIPS (inhibiting their ability to smile) and participants' electromyelogram (EMG) was monitored as a manipulation check. EEG was recorded as eighteen healthy adults (10 female) read 152 sentences about either (76) pleasant or (76) unpleasant events, along with an equal number of filler sentences. The valence of sentences was ambiguous until the occurrence of a critical target word occurring three words from the end of the sentence (e.g. "After climbing into bed, she heard her new kitten (purring/choking) close by.") Target words were approximately matched for cloze probability (6% for unpleasant, 9% for pleasant). ERPs were time locked to the pre-target word (*kitten*) that was the same in both versions of the sentence, the target word (either *purring* or *choking*) that determined the affective valence of the sentence, and the sentence final word (*by*). If facial expression impacts language comprehension, larger amplitude N400 is predicted for words read in the noisy facial feedback associated with the experimental condition.

Mean amplitude of ERPs were measured between 300-500 ms (N400) and 500-800 ms (LPC) post-word onset. ERPs to target words differed as a function of valence 500-800 ms ($F(1,17) = 8.128$, $p < 0.05$). ERPs to unpleasant targets (*choking*) were more positive than their pleasant counterparts (*purring*). Similar effects have previously been reported for ERPs to emotional language, attributed to the tendency to attend more to negative events (Kissler,J.,Herbert,C.,Winkler,I.,Junghofer,M. (2009). *Biological Psychology*, **80**, 75–83). Facial expression effects were not observed until sentence final words. The omnibus ANOVA for the sentence final words revealed no effects of Expression or Valence in the N400 interval, but revealed a significant interaction between Valence and Expression in the LPC interval, $F(1,17) = 5.360$, $p < 0.05$. Preplanned analyses revealed that the last word of pleasant sentences elicited significantly larger (more negative) N400 in the experimental (TEETH and LIPS) than in the control (LIPS ONLY) condition ($F(1,17) = 4.707$, $p < 0.05$), and ERPs in the experimental condition remained more negative in the LPC interval (Expression x Laterality x Electrodes, $F(4,68) = 3.259$, $p < 0.05$). By contrast, analysis of the unpleasant sentences revealed more positive) LPC in the experimental than the control condition (Expression x Ant-Posterior x Electrodes $F(2,34) = 4.612$, $p < 0.05$).

These data suggest emotional language gradually prompts affective responses in the body, and, these bodily responses can impact semantic retrieval operations associated with the N400. Effects were, however, confined to sentence final words, and the most robust effects of facial expressions were on the LPC, suggesting they impact continued emotional processing and decision processes induced by the valence judgment task.

Shifting viewpoints: Free indirect discourse and sensitivity to perspective-taking

Elsi Kaiser, Alexa Cohen & Emily Fedele (University of Southern California)
emkaiser@usc.edu

Tracking different perspectives is an important but controversial aspect of language comprehension. We investigated free indirect discourse (FID), which hinges on the ability to recognize perspective shift. FID presents characters' speech/thoughts without explicit quotes or embedded clauses (ex.1b,2b). How do readers recognize FID? Narratologists have identified various cues to FID (Fludernik'93, Bortolussi/Dixon'03). We investigated two cues: (i) evaluative adjectives (1b); (ii) adverbials of doubt/possibility/certainty (2b, 'possibility adverbials'). Our **first aim** is to test whether these cues have reliable effects on comprehension: Can readers spontaneously recognize shifts in perspective from the narrator to a character? Our **second aim** is to explore the cognitive mechanisms involved. We chose evaluative adjectives and possibility-adverbials, because (i) evaluative adjectives involve emotional/evaluative judgments (poor girl, stupid guy), encouraging readers to put themselves in characters' shoes/assume their perspective (cf. emotional mirroring/empathy), while (ii) possibility-adverbials involve levels of certainty/knowledge-states (probably, perhaps). Thus, their processing may differ. Specifically, we predict that perspective-taking abilities should correlate with sensitivity to evaluative adjective cues, but not possibility adverbials.

Exp.1. People read sentences with/without FID-cues (ex.(1,2); 16 targets, 30 fillers), designed so that processing the FID cue affects pronoun interpretation in sentence2. In plain sentences (1a,2a), the pronoun is ambiguous, but in sentences with FID triggers (1b,2b), if people recognize the perspective-shift, they should interpret pronouns as object-referring. After reading each sentence, participants (n=36, 19 female) indicated, on 6-point scale, who the pronoun refers to (e.g. "Who was sick?" Mary 1 2 3 4 5 6 Kate). Afterwards, participants completed the Perspective Taking/Spatial Orientation Test (Hegarty/Waller'04).

(1a) Mary looked woefully at Kate. She was sick.   (she=ambiguous)
(1b) Mary looked woefully at Kate. Poor girl; she was sick.
(2a) Luke glanced at Tom warily. He'd put toothpaste in the shampoo bottle again.   (he=ambiguous)
(2a) Luke glanced at Tom warily. He'd probably put toothpaste in the shampoo bottle again.

Results. Both evaluative adjectives and possibility-adverbials resulted in pronouns being more likely to be interpreted as referring to objects, compared to pronouns in 'plain' sentences (p's<.05)—subtle cues to FID have clear effects on pronoun resolution. To explore perspective-taking and FID, FID Sensitivity Scores were calculated for each participant (difference between their average plain-sentence rating and average FID-adverbial/FID-adjective-sentence rating). We find that (i) sensitivity to evaluative adjectives correlates with spatial-perspective abilities (p<.05), but (ii) there is no correlation between sensitivity to possibility adverbials and spatial perspective-taking performance (p>.7).

Exp. 2. To verify that the FID cue sensitivity observed in Exp.1 was indeed due to people recognizing a shift in perspective, in Exp.2 participants (n=24) first completed the same rating task as in Exp.1, and after having rated all items, were *asked to explicitly indicate whose voice or point-of-view/opinion is expressed in the second sentence* (subject/object/narrator). The rating results replicate Exp1: Both FID cues have clear effects on pronoun interpretation (p's<.01). Moreover, the viewpoint task elicited more 'subject-character viewpoint' responses for sentences with both types of FID cues (p's<.01) and more 'narrator viewpoint' responses (p<.01) for plain sentences: People can successfully recognize both evaluative adjectives and possibility adverbials as signaling a shift from the narrator to a character's perspective.

In sum, Exp1-2 showed that seemingly subtle cues such as evaluative adjectives and possibility adverbials have clear effects on discourse processing (pronoun resolution): Readers spontaneously recognize shifts in perspective from narrator to character, as confirmed by Exp.2. Furthermore, Exp.1 showed that only people's sensitivity to evaluative adjectives is reliably correlated with spatial perspective-taking, suggesting that different FID triggers may rely on different cognitive mechanisms.

Visuospatial grouping influences expectations about upcoming discourse
Elsi Kaiser & David Cheng-Huan Li (University of Southern California)
emkaiser@usc.edu

Understanding the relations between events and entities is an important aspect of discourse-level linguistic processing (e.g. Kehler et al.,'08). How we conceptualize and recognize events is also of interest to cognitive psychologists: When we observe the world, we are faced with a continuous 'stream of activity' that we segment into meaningful events (Zacks/Swallow'07). We explored whether and how *visuospatial cues*–e.g. those provided by panels/frames in comics–contribute to our understanding of narrative representation and event segmentation. Panel segmentation is a key aspect of comics (Eisner 1985/2008, Cohn'12).

We conducted a psycholinguistic experiment investigating how segmenting visually-presented events into panels in different ways–in other ways, changing their visuo-spatial encoding–influences readers' conceptualization of narratives. In particular, we wanted to explore whether and how differences in framing / visuospatial grouping influence what aspects of the event are regarded as more important/cognitively prominent, as measured by people's expectations about upcoming discourse.

We tested whether depicting a transitive event in two panels or one panel (examples below) influences readers' mental representations, by means of a fragment-continuation task: The closing panel shows two new characters, one of whom says "…and then…", and participants wrote a continuation for this fragment. Fragment continuation can probe people's expectations about upcoming discourse and measure how prominent/salient/important different aspects of the event are in people's minds. Do people write about **(i)** the pictured event (e.g., *…and then Lisa tripped Mary*), **(ii)** a consequence resulting from the event (e.g., *…and then the girl sprained her wrist*), **(iii)** subsequent events that are not causally related (e.g. *…and then the mean girl also tripped Andy*) or something else? (Names of characters were not prespecified, participants could refer to them however they wished).
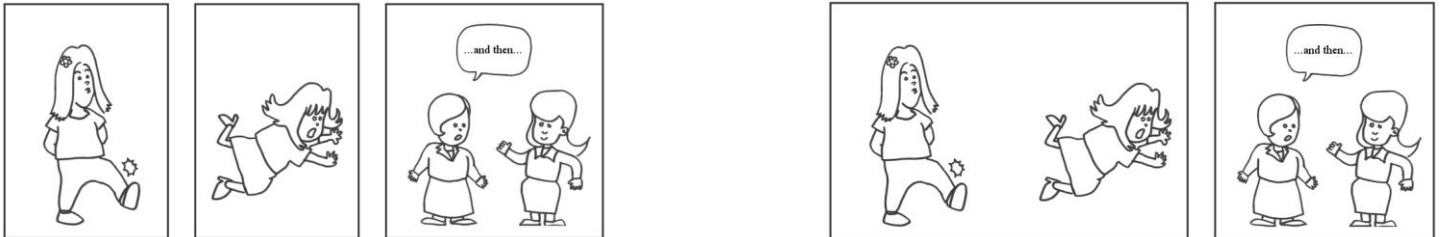
Design. Twenty-eight native English speakers participated. Each person saw 24 comics (8 targets, 16 fillers). To control semantics, the eight critical strips depicted active events with agents and patients (e.g., *tripping, tickling, punching)*. The only difference between one-panel/two-panel versions was the 'gutter' between the characters in the two-panel version; physical distance was unchanged. We manipulated the visuospatial properties of the critical event by grouping both the agent and patient in one frame/panel or by separating them into two panels. (Two-panel versions had "subject-to-subject" transitions, in McCloud'93's terminology)

Results. Two-panel sequences resulted in significantly more *result/consequence-type completions* than one-panel sequences (two-panel = 47% vs. one-panel = 30%, p's<.02): When the visual information was split into two panels, participants were more likely to focus on the result/consequences of the event than in one-panel versions. In contrast, one-panel versions resulted in significantly more *continuations describing the pictured event* than two-panel versions (two-panel = 28% vs. one-panel = 43%, p's<.02).

Discussion. Visuospatial grouping/segmentation guides expectations about discourse: Isolating the patient in its own panel makes the consequences of the event (even when not explicitly shown) more prominent / salient (more likely to be mentioned). In contrast, depicting agent and patient together emphasizes their interaction (the pictured event). Psycholinguistic models of discourse need to include information about event segmentation, in addition to known factors such as coherence relations, verb semantics and referent salience.

Evidence for a rational probabilistic account of Gricean implicatures
Daniel Grodner (Swarthmore College) & Benjamin Russell (Brown University)
dgrodne1@swarthmore.edu

The standard account of how perceivers recover implicit content is via a process of rational psychosocial inference (Grice, 1975). By assuming that speakers are cooperative, perceivers can generate explanations for why a speaker chose to utter a given expression over plausible alternatives. For instance, uttering the weak expression ω often implicates (1a). A cooperative speaker should make the strongest assertion compatible with their knowledge. Thus the perceiver can infer that the speaker doesn't believe the stronger statement σ as in (1b). If the speaker is also thought to have an opinion about whether the stronger alternative is true ($B_{spk}\sigma$ or $B_{spk}(\text{NOT}\sigma)$), this inference will be strengthened to (1c), which, for a trustworthy speaker, is equivalent to (1a) The Gricean account has recently been challenged on the grounds that it cannot account for so-called local implicatures (Chierchia et al., 2012; Gajewski & Sharvit, 2011). Gricean reasoning predicts that (2) implies the global implicature (2a). However, intuitively (2) yields the local implicature in (2b). To account for this, some investigators have proposed that these implicatures are computed in the grammar by optionally inserting a silent operator equivalent to *only* before *some.* If this is right, the rational appearance of many implicatures is an illusion.

The present study tests the predictions of extending the standard Gricean account in two ways, each of which follows from principles of rational inference, to account for local implicatures. First, a rational perceiver should impute opinionatedness not just to speakers, but to other agents such as sentential subjects (van Rooij and Schulz, 2004). This can generate the implicature in (2b) without positing a grammatical operator (2c). Second, a rational perceiver should commit to the local implicature probabilistically, using a Bayesian combination of the evidence for the global implicature and the prior likelihood that the subject is opinionated (Russell, 2012). Opinionatedness is a priori less likely for embedding predicates that convey stronger attitudes: For any σ, Jim is less likely to be certain σ or certain NOTσ than he is to think σ or think NOTσ. Thus the likelihood of generating the local implicature should decrease for stronger predicates.

To evaluate the extended model, 120 participants were asked to judge the likelihood of the local or global implicature for sentences that contained *some* in the embedded clause of a weak (e.g., *think*) or strong (e.g., *be certain*) attitude. 16 stimuli like (3) were intermixed with 50 fillers. Mixed effect modeling revealed a reliable interaction between predicate strength and implicature type ($z=3.1$, $p < .01$). Weak predicates were significantly more likely to generate a local implicature than strong ones, but there was no effect of predicate strength for global implicatures. To establish whether the effect of opinionatedness was graded or binary, a separate set of 120 participants judged the prior probability of the local and global queries without the trigger sentence. For each item, opinionatedness was calculated as the sum of the mean rating for the two queries. The rational account models the likelihood of the local implicature as the product of the evidence for the global implicature and opinionatedness. The model prediction was strongly correlated with the data ($r=.54$, $p<.01$).

This work supports a rational model of implicature generation. Unlike a grammatical account, this model both predicts how the strength of the local implicature reading will vary across contexts and situates implicature phenomenon in a broader theory of psychosocial reasoning.

ω = Some voles hibernate $\qquad\qquad$ σ = All voles hibernate
(1a) Not all voles hibernate $\qquad\qquad\qquad\qquad\qquad$ (1b) NOT($B_{spk}\sigma$) $\qquad\qquad$ (1c) $B_{spk}(\text{NOT}\sigma)$
(2) Jim thinks some voles hibernate (= $B_{Jim}\omega$) $\qquad$ (2a) $B_{spk}(\text{NOT}(B_{Jim}\sigma))$ $\qquad$ (2b) $B_{spk}(B_{Jim}(\text{NOT}\sigma))$
(2c) $B_{spk}(\text{NOT}(B_{Jim}\sigma)) + B_{spk}(B_{Jim}\sigma$ or $B_{Jim}(\text{NOT}\sigma))$ = $B_{spk}(B_{Jim}(\text{NOT}\sigma))$
(3) **Context:** Mary and Jim were studying for a big biology midterm. **Trigger:** Jim (thought/was certain) that some voles hibernate.
$\qquad$ **Global query:** Jim (thought/was certain) that all voles hibernate.
$\qquad$ **Local query:** Jim (thought/was certain) that not all voles hibernate.

Sensitivity to local discourse vs. global communicative context in gradable adjectives

Christina Kim, Andrea Beltrama (University of Chicago), Kristen Syrett (Rutgers University), Ming Xiang & Chris Kennedy (University of Chicago)
cskim@uchicago.edu

The interpretations of gradable adjectives (GAs) relies on contextually determined standards: the truth of (1) can only be evaluated after a standard for what counts as *tall* is fixed (Kennedy and McNally 2005, Syrett et al. 2009). This study tests the hypothesis that two subclasses of GAs are sensitive to different aspects of the context in fixing their standards: *Relative GAs* (*tall*) use open scales, with the discourse context supplying a standard of comparison based on a comparison class. *Absolute GAs* like (2) have closed scales, with standards fixed to maximal (*full*) or minimal (*wet*) degrees. However, although the standards of comparison are fixed, context nevertheless determines a *standard of precision*—to evaluate (2), one must determine how precise one can be given a communicative situation [Lasersohn 1999]. In three experiments, we show that standards of comparison for Relative GAs are sensitive to the local discourse context, while standards of precision for Absolute GAs are sensitive to properties of the broader communicative context.

Experiments1-2 (n1=28,n2=20) showed that Relative but not Absolute GAs depend on the local discourse context. We created sets of images representing six points on a continuum characterized by an adjective pair (*tall-short glass*). Participants indicated for each image and adjective pair whether they considered the pictured object to be e.g. *tall*, *short*, or *neither*. In addition, they either saw items individually, interspersed with trials with other adjective-noun pairings so that each trial represented its own local discourse context (Exp1), or with the images associated with the same adjective-noun continuum presented as a group (Exp2), thereby providing a comparison class in context. Only for Relative GAs, we found: (i) individual presentation increased dependence on item-specific prototypes: the identity of the head noun accounted for 75.4% of the random effects variance in a mixed effects regression model, as compared to 67.4% when items were presented in groups; and (ii) the position of an item on an adjective-noun continuum was a reliable predictor of response for grouped (β=.40, p<.0001), but not for individual presentation (β =-.05, p>.1).

Experiment 3 (n=79) showed that Absolute but not Relative GAs are sensitive to the global context. Participants saw pairs of objects (an almost full and a perfectly full glass (Absolute); a tall and a taller glass (Relative)), and responded to auditorily presented instructions like (3 and 5). On half of the Absolute GA trials, the fuller of the two glasses was perfectly full (Felicitous and precise); in the other half, the fuller glass was not quite perfectly full (Imprecise or Infelicitous). On half of the Relative GA trials, the taller of the two glasses was from the high end of the continuum, and from the lower end in the other half. We manipulated whether an instance of an adjective-noun pair ({*full,tall*} *glass*) was the first occurrence of that adjective in the experimental session, or whether it was preceded (at a 4-13 trial lag) by a prior occurrence of that adjective. Of interest was whether setting a standard of precision for an Absolute GA on a prior trial would influence the standard used to evaluate a subsequent instance of the same adjective (3-4); and similarly, whether setting a standard of comparison for a Relative GA on a prior trial influences the standard use on a subsequent trial with the same adjective (5-6). Absolute GAs showed a clear ordering effect: Infelicitous (i.e. imprecise) trials were rejected more often after a Felicitous trial than when there was no prior trial that required setting a standard of precision (β =-2.15, p<.0001). Relative GAs, by contrast, show no ordering effects (β =.51, p>.1).

Conclusion. These results provide additional experimental evidence for a distinction between Relative and Absolute GAs; further, we show that these classes of expressions can be differentiated with respect to their sensitivity to different aspects of the context in a way that suggests that standards of comparison are more tightly linked to the discourse representation, while standards of precision are more closely tied to broader features of the context.
(1) The glass is *tall.* (2) The glass is *full.*
(3) *Choose the full glass.* (4) [Felicitous before Infelicitous] or [Infelicitous before Felicitous]
(5) *Choose the tall glass.* (6) [High end of scale before Low end] or [Low end before High end]

Objects and actions in dis-agreement
Jason Schoenberg & Heidi Lorimor (Bucknell University)
hml003@bucknell.edu

Thirty years of research on agreement attraction has demonstrated that subject-verb agreement is determined by several factors, including the grammatical number of the nouns involved, the notional number value of the referent, and the morphophonological properties of the elements within the subject noun phrase (e.g., Badecker & Kuminiak, 2007; Bock & Middleton, 2011; Haskell & MacDonald, 2003). To date, research has primarily focused on noun phrases that depict people, objects, or groups, (e.g., "the label on the bottles", "the family of rats"), as those are easily quantifiable and manipulable in experimental paradigms. However, no one has taken a systematic look at how subject-verb agreement works when the subjects of sentences are actions. This is surprising, as speakers frequently talk about actions like "writing an abstract" or "grading papers".

We created 16 sets of sentence preambles to investigate whether differences in semantic readings (whether a phrase is interpreted as an action or an object) would influence the rate of agreement attraction. We also manipulated the morphological properties of the head noun (i.e., how "verbal" the subject noun was) to see whether the morphological properties of the subject head nouns would influence agreement patterns. For each of the sets of preambles, we also created 2 videos (one depicting an object interpretation, the other depicting an action interpretation) that participants saw as they heard the preamble. Crucially, for the deverbal condition (in which the noun is morphologically derived from a verb), the exact same preamble was used to accompany the object and action videos, so any differences in agreement patterns between the two conditions had to come from the difference in semantic interpretation, and not from the words themselves. We collected norming data for the preambles on imageability and naturalness, and for the preambles and their videos on notional number.

In a sentence completion task, 48 native English speakers listened to preambles and watched the corresponding videos. Their instructions were to repeat the preambles and complete the sentences. Sample items and the proportion of plural verbs produced in each condition are below.

Type of head noun	Corresponding video	Sample item	Proportion plural verbs
Simple noun	object	The bouquet of flowers	0.24
Deverbal noun	object / action	The arrangement of flowers	0.10 / 0.05
Gerund	action	Arranging flowers	0.06

Logistic mixed-effect regression models showed that notional number played a significant role in the types of agreement produced, as items that were more notionally plural were more likely to take plural agreement. In addition, the factors of semantic reading (objects vs. action) and morphological derivation (deverbals & gerunds vs. simple nouns) were also significant, with less agreement attraction occurring in the action readings and when the nouns were morphologically derived from verbs. None of the other factors were significant. This demonstrates that research on subject-verb agreement needs to be expanded to consider how agreement is computed when the subject of the sentence is an action, and also suggests that the morphological make-up of a noun itself may prove to be an important factor in agreement production.

References
Badecker, W. & Kuminiak, F. (2007). Morphology, working memory, and agreement in sentence production: Evidence from gender and case in Slovak. *Journal of Memory and Language, 56(1)*, 65-85.
Bock, J.K., & Middleton, E.L. (2011). Reaching agreement. *Natural Language & Linguistic Theory, 29(4)*, 1033-1069.
Haskell, T.R. & MacDonald, M.C. (2003). Conflicting cues and competition in subject-verb agreement. *Journal of Memory and Language, 48(4)*, 760-778.

Filling in the blanks in morphological productivity: A word-completion task
Kyle Mahowald, Timothy O'Donnell, & Joshua B. Tenenbaum (MIT)
kylemaho@mit.edu

People have strong intuitions about the relative probabilities of linguistic structures such as words, morphemes, or word combinations, and these intuitions often systematically differ from empirical corpus frequencies (see, e.g., Smith & Levy, 2011). In one classic study (Tversky & Kahneman, 1973), participants judged partial word patterns including an English suffix, such as _ _ _ _ i n g, to be consistent with a greater number of English words than patterns that did not include full suffixes, like _ _ _ _ _ n _, despite the fact that every word that matches the former pattern also necessarily matches the latter. TK interpreted this fallacy as a failure of rational probabilistic judgment— as a sign that people make judgments based on heuristics such as the greater availability of linguistic representations in memory (with identifiable morphemes like –*ing* aiding retrieval of words that include them). While these results clearly show that people access words via morphemes more easily than via single letters, we propose that this is hardly a failure of probabilistic inference. On the contrary, it follows from a fundamental design feature of the linguistic system – productivity, the ability to produce and comprehend novel linguistic forms – and it provides a tool for studying sophisticated mechanisms of probabilistic inference at the heart of human linguistic processing. Because –*ing* is a highly productive morpheme, there are many possible English words that can be derived using it, more possible words than will be observed in any finite corpus sample, or, indeed, in the speakers own linguistic experience. The productivity of –*ing* causes participants to "hallucinate" possible but hitherto unobserved word forms. Here, we test this hypothesis by extending T&K's work using a variety of English derivational suffixes of varying levels of productivity. Using a rational Bayesian model of productivity and storage known as Fragment Grammars (FG; O'Donnell, 2011), we derive estimates of the frequencies for both observed and novel (out-of-sample) words. The measure accounts for frequent existing words as well as novel, productively-derived words. We predict that productive suffixes (e.g., -*ness*) are more likely to have their frequencies overestimated than lower productivity suffixes (e.g., -*ity*).

Methods. Using Mechanical Turk, we presented 206 participants with 105 word frames (like _ _ r _). Participants estimated how many times a given pattern would occur in a 100,000 word book. The word frames fell into one of these categories: (i) full-suffix frames like _ _ _ _ n e s s, (ii) partial-suffix frames like _ _ _ n _ s _, and (iii) frames based on mono-morphemic words like r _ _ d (*road*, *reed*, etc.). We ultimately analyzed 40 suffixes of 3 or more letters drawn from O'Donnell (2011). For suffix-derived frames, each participant saw at most one instance of a given suffix. Partial-suffix frames were created by randomly deleting letters from full suffix frames. Mono-morphemic frames were sampled to have a wide spread of frames, from those with many possible completions like s _ _ _ _ to those with few, like b r i c _.

Overall results. A linear mixed-effect model of participant guesses was fit with the following centered predictors: (i) number of words tokens consistent with the frame (in some sense the "correct" answer) from SUBTLEX, (ii) number of word types consistent with the frame, (iii) the number of letters missing from the frame, (iv) the number of letters present, (v) the interaction between missing and present letters, and (vi) the frame type—all of which were significant predictors. There was a maximal random effect structure for participant. As predicted by K&T, full suffix frames were inflated relative to a baseline for partial suffix frames while controlling for frequency (β=.97, t =20.43, $\chi^2(1)$ =61, pχ^2< .001), and both full and partial frames were overestimated relative to mono-morphemic frames (β= .35, t =8.41, $\chi^2(1)$=229, pχ^2 < .001).

Full suffix results. To further investigate the role of morphological productivity for the full suffix frames, we used predictions from FG that represent estimates for the total probability of all existing and potential words that use a given suffix. We fit a mixed effect model with a fixed effect of this quantity, the same controls as above, and a maximal random effect structure for subject and a random intercept for suffix with a slope for stem length. A higher marginal probability as estimated by FG was predictive of a higher estimate in the task (β=.09, t=2.35, $\chi^2(1)$=10.5, pχ^2 < .01).

Conclusion. The role of morphological productivity in these estimates suggests that K&T's effect is likely caused by a "hallucinatory" effect of productive morphemes. Because productive morphemes can give rise to unbounded numbers of novel forms, the presence of a productive suffix in isolation causes overestimation of the frequency of the pattern, with more productive morphemes leading to greater rates of overestimation.

How different levels of syntactic flexibility influence language production in Mandarin

Xin Zhao, Elsi Kaiser (University of Southern California)
xinz@usc.edu

During grammatical encoding, speakers select lexical items and assemble them into a syntactic structure (Levelt'89; Bock/Levelt'94). Existing work (Ferreira'96) supports an incremental model of grammatical encoding: Having two available syntactic structures (flexible condition) makes production easier (less errors, shorter latencies) than having only one available structure (non-flexible condition). However, it is not yet known whether effects of syntactic flexibility are _categorical_ (all that matters is the presence/absence of flexibility) or _gradient_, such that _higher levels of flexibility_ (e.g. 3-4 choices) could have even stronger facilitatory effects.

We conducted a **production study** exploring how different levels of syntactic flexibility influence production in Mandarin Chinese. When Mandarin locative resultatives occur with the preposition '_ba'_, only one syntactic structure is possible (ex.1a), regardless of whether the theme argument is indefinite (e.g. _one-CL hat,_ 'a hat') or definite (_that-CL hat,_ 'that/the hat'). However, locative resultatives without '_ba'_ result in greater syntactic flexibility: when the theme is indefinite (_one-CL hat_), four structures are available (1a-d). We refer to this as quadruple flexibility. When the theme is definite (_that-CL hat_), three structures are available (1a-c), i.e., ternary flexibility. Crucially, all sentences in (1a-d) have the same basic meaning: _he will put a/the hat on the chair_. Participants (n=20) were asked to construct one sentence at a time using all words shown on a computer screen. We manipulated (i) the theme's definiteness and (ii) w/ or w/o _ba_ to create four conditions (see table).

(1) a. ta yao **ba** {**yi-ding/na-ding**} maozi da-zai yizi-shang	[Locative _ba_-construction]
he will **BA** {**one-CL/that-CL**} hat put-at chair-up	
b. ta yao da {**yi-ding/na-ding**} maozi zai yizi-shang	[Locative dative]
he will put {**one-CL/that-CL**} hat at chair-up	
c. ta yao zai yizi-shang da {**yi-ding/na-ding**} maozi	[Preverbal-locative resultative]
he will at chair-up put {**one-CL/that-CL**} hat	
d. ta yao da-zai yizi-shang {**yi-ding/*na-ding**} maozi	[Locative double object construction
he will at chair-up put {**one-CL/ *that-CL**} hat	ungrammatical w/ def theme]

Results. Compared to non-flexible conditions, both ternary and quadruple flexibility conditions resulted in significantly _lower_ error rates ($p<.05$), _higher_ correct response rates (p's<.05), and _shorter_ production latencies ($p<.05$). But ternary- and quadruple-flexibility conditions did not differ from each other on these measures, indicating that higher levels of flexibility do _not_ result in further facilitation. Our experiment represents the first investigation of different levels of syntactic flexibility, and _suggests its facilitatory effects are categorical, not gradient,_ providing new insights into the nature of incremental production.

	Preposition BA absent	_Prep. BA present_
Indef Theme	**quadruple flexibility**	**non-flexibility**
	1) locative _ba_-construction; 2) locative dative; 3) preverbal-locative resultative; 4) locative double object.	1) locative _ba_-construction.
Def. Theme	**ternary flexibility**	**non- flexibility**
	1) locative _ba_-construction; 2) locative dative; 3) preverbal-locative resultative.	1) locative _ba_-construction.

Implicit naming in the visual world paradigm

Daniel Pontillo, Anne Pier Salverda & Michael Tanenhaus (University of Rochester)
dpontillo@bcs.rochester.edu

We present a theoretical perspective on the vision-language interface in which visual representations are automatically activated for spoken words and utterances, whereas the retrieval of phonological representations associated with pictures and scenes is under strategic control. This view, which is supported by the absence of reverse Stroop effects and by the literature on linguistic relativity, predicts that (1) linguistically mediated fixations are primarily driven by visual/perceptual (and conceptual) representations evoked by a spoken word and (2) implicit naming of objects or pictures in visual world studies will only occur when name retrieval facilitates task performance.

To directly test the visual/perceptual and implicit naming hypotheses, we developed a set of 14 pictures with synonymous names (e.g., couch/sofa) and cohorts associated with each name (cow and soda, respectively). Norms established that (1) when the picture was followed by a name, both names were rated as equally good and (2) when shown the picture and asked to generate a name, one name was strongly dominant. The visual/perceptual hypothesis predicts that dominance will affect fixations only when aspects of the experimental task encourage the participant to retrieve the picture's name.

In Experiment 1, we modified the standard setup of a four-picture spoken word recognition study in the visual world paradigm by masking a picture to create memory demands, which are known to promote linguistic encoding. Four pictures were displayed in a grid, and one of the pictures was highlighted by a red square. After 1.5 seconds, the highlighted picture was covered with a semi-opaque mask. The position of the picture with the mask was randomized, and trials were counterbalanced for whether the target was the masked picture. Additional filler trials controlled for other contingencies. In the test trials, the target was one of the two cohorts, while the synonym picture was masked. There were significantly more looks to the masked picture (e.g. couch/sofa) when the target word was a cohort of the dominant synonym name (cow) than when the target word was a cohort of the subordinate synonym name (soda). This result suggests that participants retrieved a name associated with the masked synonym picture. Thus, robust name typicality effects emerge under conditions in which memory demands encourage linguistic encoding.

Experiment 2 presented the same stimuli with the same preview, but no pictures were masked. The visual/perceptual hypothesis predicts that name typicality effects should be eliminated under these conditions (which are typical for visual world experiments). The results supported this prediction: We found strong cohort effects, with a similar proportion of looks to the synonym picture regardless of whether the target was a cohort of the dominant or subordinate name associated with the synonym picture.

Taken together, these results strongly support the visual/perceptual mapping hypothesis, and they suggest that the activation of phonological codes is strategic, adopted only to accommodate to memory demands.

Theory of mind drives efficient language production

Peter Graff, Zoe Snape (MIT), Jeremy Hartman (UMass Amherst) & Edward Gibson (MIT)
graff@mit.edu

Introduction. Recent work has shown that speakers produce language in communicatively efficient ways (e.g, Jurafsky et al. 2001, Aylett & Turk 2004, Levy & Jaeger 2007). For example, Jaeger (2010) presents a corpus study of complementizer *that*-mentioning (e.g., *I said (that) John would come to the movies.*) showing that speakers preferentially mention *that* if the matrix verb is biased against appearing with a complement clause (CC-bias), thus keeping the information density of their utterances uniform over time. However, whether such behavior derives from speaker-internal or listener-oriented factors is still largely unknown (see Gahl et al. 2012 for discussion). In this paper we present evidence that the efficiency of a speaker's production choices with respect to *that*-mentioning depends at least in part on her theory of mind. We argue that this result constitutes preliminary evidence in favor of listener-oriented accounts of efficiency in language usage (e.g., Clark & Brennan 1991).

Methods. We extracted 10 frequent verbs with varying CC-biases from Jaeger (2010; *find, think, read, say, mean, know, see, feel, hear, guess*). Next, we generated 20 sentence frames that plausibly allowed for any of those ten verbs to occur as the matrix verb (e.g., *Bobby ____ (that) your artwork was excellent.*). Verbs were assigned to sentence frames in a Latin square design, such that every verb appeared in 2 different frames for each participant. Participants were asked to choose between two variants of a given sentence differing only with respect to the presence of the complementizer *that*. We additionally included 20 filler items, whose acceptability varied strongly based on the presence of that (e.g. *I asked (that) what you were doing*, *he loves it (that) I wear sweaters*). Before completing this task, participants were asked take the "Reading the Mind in the Eyes" test for theory of mind (revised version; Baron-Cohen 2001; RMITE). To allow us to identify participants who provided random responses to the RMITE test, participants were asked answer eight randomly selected RMITE questions a second time after completing the language task.
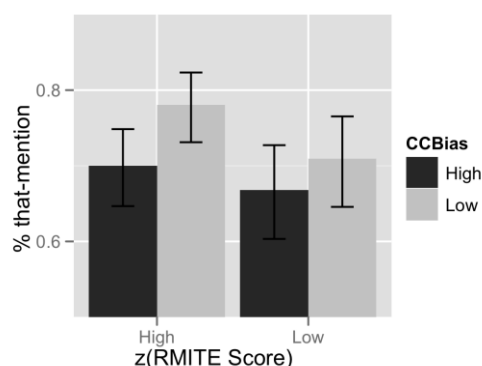
Results. Eighty workers participated in the experiment through Amazon's Mechanical Turk crowdsourcing platform. Participants who did not respond consistently to at least 6 out of the 8 RMITE repeat questions and/or did not provide correct judgments on at least 70% of the filler items were excluded, leaving data from 53 participants for analysis (higher consistency/filler thresholds give analogous results). We fit a mixed logit model predicting that-mentioning from the z-transformed RMITE score of the participant, the log-transformed complement clause bias of the verb (verb-based; Jaeger, 2010; CC-bias) and their interaction. The model included random intercepts for participant and sentence frame, random slopes for all fixed effects grouped by sentence frame, and a random slope for CC-bias grouped by participant. We find a highly significant effect of CC-bias on that-mentioning (β=-0.33, z=-5.44, p<.00001), thus replicating Jaeger's (2010) corpus result for the first time in an experiment. We find no main effect of RMITE score (β=0.06, z=0.36, p=0.72), however, we do find the predicted interaction of RMITE score and CC-bias (β=-0.14, z=-2.27, p<.03), such that the effect of CC-bias on that-mentioning is smaller for speakers with lower RMITE scores (see Figure).

Discussion. Our results suggest that efficient production choices such as those observed by Jaeger (2010) depend at least partially on the speaker's ability to emphasize with other human beings. This is not expected if efficient production choices are driven entirely by properties of the language production system, which have so far not been shown to interact with a speaker's theory of mind.

Individual differences in reading styles and the use of implicit causality as a pronoun resolution cue

Arnout Koornneef & Ted Sanders (Utrecht University)
a.w.koornneef@uu.nl

Rayner et al. (2009) suggested that readers that make long saccades and frequently look back in a text adopt a *risky reading* strategy: they 'guess' how a sentence or text will continue. In the present study we explored this hypothesis by presenting texts that confirmed or disconfirmed verb-based implicit causality (IC) expectations about who would be talked about next (e.g. that "*David apologized to Linda because he/she*..." would continue about David, not Linda).

Forty-six native speakers of Dutch read short stories containing IC-consistent and IC-inconsistent pronouns while their eye-movements were monitored (see ex. 1).

1. Subject-biased verb, introduction
    David and Linda were both driving pretty fast. At a busy intersection they crashed hard into each other.
    a. *Consistent pronoun*
        David apologized to Linda. He according to the witnesses was the one to blame.
    b. Inconsistent pronoun
        David apologized to Linda. She according to the witnesses was not the one to blame.

The eye-movement recordings were applied in two ways. First of all, we computed commonly reported reading time measures (e.g. first-pass duration, regression-path durations etc.) and in addition, we used the eye-tracking data to determine the reading style of the participants. Based on Rayner's proposal that risky readers make longer saccades and look back more often in a text, we used a median split procedure for these two dimensions to assign the participants to four groups: (i) risky readers (long saccades, many regressions), (ii) conservative readers (short saccades, few regressions), (iii) fast readers (long saccades, few regressions), (iv) slow readers (short saccades, many regressions).

Our main finding was that although both risky readers and conservative readers were sensitive to IC information (as indicated by reading time delays in the inconsistent condition), a difference was observed with respect to *when* IC effects emerged. Risky readers were using IC right at the moment they encountered the pronoun, yet a three-word delay was observed for the conservative readers. Furthermore, we did not observe effects of IC for the fast and slow reading groups. Instead these readers preferred to connect the pronoun to the first mentioned entity (i.e. the subject).

Our results are consistent with the idea that people that make relatively long saccades and many regressions are using a more risky (or proactive) processing strategy than readers making shorter saccades and less regressions. Moreover, whether implicit causality is activated and used at all seems to depend on the characteristics of the reader. As such our results complement studies that have shown that only skilled readers activate IC (Long & DeLey, 2000) and that early IC effects typically emerge when readers are primed to use a proactive, (i.e. top-down, knowledge driven) processing strategy (De Goede et al, 2010).

References

De Goede et al. (2009) The effect of mood on anticipation in language comprehension: An ERP study. *Presented at the 22nd Annual meeting of the CUNY Conference on Human Sentence Processing*, Davis, CA.

Long, D. & De Ley, L. (2000). Implicit causality and discourse focus: The interaction of text and reader characteristics in pronoun resolution. *Journal of Memory and Language 42(4),* 545-570.

Rayner, K., Castelhano, M. & Yang, J. (2009). Eye movements and the perceptual span in older and younger readers. *Psychology and Aging, 24*, 755-760.

Case licensing in processing: Evidence from German
Shayne Sloggett (University of Massachusetts Amherst)
sslogget@linguist.umass.edu

Grammatical case has frequently been advanced as a heuristic for structuring an incoming parse: case provides cues to grammatical role [1], alleviates the cost of scrambled structures [2], and affects the ability of arguments to participate in agreement [3]. Less work, however, has looked at case licensing in processing—the parameters which control whether or not a case-marked element is perceived as illicit. Previous work has shown that erroneous dative case is more readily rejected than erroneous accusative [4] and that case features compete for representational space [5]. However, this work leaves open the question of how case-marked elements are linked by the sentence processor to their licensing environments. In the present study I give evidence that dative case can be used as a cue for memory retrieval, and propose that case licensing is achieved via feature matching in working memory [6]. This hypothesis ties case licensing to previous work on the processing of agreement dependencies, where cue-based retrieval has been advanced as an explanation for the spurious licensing of plural agreement morphology [7]. If case licensing is implemented in the same architecture, we predict that case features should be similarly prone to illusory licensing effects.

To test this prediction, I constructed 20 sentences manipulating the factors *grammaticality* and *distractor*. These factors varied the case of a dependency target (DAT: grammatical; ACC: ungrammatical) and a distractor noun (DAT: distractor; ACC: no distractor) respectively in a 2×2 design.

(1) Das Kind hat <u>seinen/*seine</u> Großeltern mit **einem** Lächeln **<u>gedankt</u>**.
   The child has <u>his.Dat/*his.Acc</u> grandparents with **a.Da**t smile **<u>thanked</u>**.
(1) Das Kind hat <u>seinen/*seine</u> Großeltern für das Geschenk <u>gedankt</u>.
   The child has <u>his.Dat/*his.Acc</u> grandparents for the gift <u>thanked</u>.

The clause-final verb idiosyncratically selected for dative case and was separated from the object by a PP containing the distractor. If case is a feature on par with morphological number, and licensing relies on retrieval, then we should observe a difference between the ungrammatical versions of (1) and (2): ungrammatical sentences with a dative-marked distractor (bolded in (1)) should be accepted more frequently than ungrammatical sentences with no distractor. Preliminary results of a speeded grammaticality judgment experiment (n=14) bear out this prediction:

	Distractor	No Distractor
Grammatical	0.93(.03)	0.86(.05)
Ungrammatical	0.50(.08)	0.69(.08)

Table 1. *Mean Accuracy (SE)*

Logistic linear mixed effects modeling reveals a reliable main effect of grammaticality (z=5.33, p<.01) and a significant interaction of grammaticality and distractor (z=2.61, p<.01). A planned comparison of the ungrammatical conditions shows a reliable difference (t(12)=-3.27, p<.01), suggesting that the ungrammatical, distractor condition is driving the interaction. This result both supports the hypothesis that case licensing is a process of feature matching in retrieval, and extends the list of grammatical illusions to a novel dependency. This latter finding is particularly interesting given the relative paucity of illusory licensing phenomena.

References. [1] Skopeteas, Fanselow, & Asatiani. 2012. *Case, word order and prominence*. [2] Bornkessel, Schlesewsky, & Friederici. 2002. *Cognition*. [3] Badecker & Kuminiak. 2007. *Journal of Memory and Language*. [4] Bader, Meng, & Bayer. 2000. *Journal of Psycholinguistic Research*. [5] Bader & Bayer. 2006. *Case and Linking in Language Comprehension*. [6] Lewis & Vasishth. 2005. *Cognitive Science*. [7] Wagers, Phillips, & Lau. 2009. *Journal of Memory and Language*.

Poster Session 3

How bizarre: Sentence processing and memory

Peter C. Gordon, Matthew W. Lowder, Miri Besken, & Neil Mulligan (University of North Carolina at Chapel Hill)
pcg@email.unc.edu

Both psycholinguists and memory researchers place great theoretical emphasis on the encoding of information conveyed by language, but their distinct goals have led them to approach this topic in very different ways. Psycholinguists study encoding using a range of techniques that provide online evidence about the processes of sentence comprehension. In contrast, memory researchers use performance in subsequent memory retrieval tasks to make inferences about encoding. This project combines these two approaches by recording eye movements during reading and interpreting this online measure both in relation to characteristics of the sentence and subsequent performance on incidental memory tasks. As shown below, the sentences varied the semantic congruity of the modifier and patient (e.g., large carrots) with the preceding instrument/action; alternative patients (e.g., yellow balloons) were used for counterbalancing[1]. This precisely localized change determines whether the meaning of the sentence is *common* or *bizarre*, a factor that affects both sentence processing and memory.

> [Common]   John used a knife to chop the large carrots for the party.
> [Bizarre]   John used a pump to inflate the large carrots for the party.

Subjects read 12 common and 12 bizarre sentences while making plausibility judgments, then received a surprise memory test. Half the subjects got a free-recall task where they were instructed to write down as much as they could remember about the sentences. The other half got a cued-recall task, where cues could consist of the modifier/patient (The large carrots were ___ with a ___) or the instrument/action (A knife was used to chop the ___ ___). Memory performance for a sentence was scored as correct if any one of the critical words was recalled accurately. Sentence reading time was longer for bizarre than common sentences (see table). For free recall, memory performance was higher for bizarre than for common sentences. This finding is consistent with previous work and seems to suggest that greater reading time for bizarre sentences corresponds to greater effort at encoding which in turn enhances later efforts at recall. However, for the cued-recall task, memory performance was lower for bizarre than for common sentences (a separate fill-in-the-blank norming study showed that the increase in cued-recall performance for common sentences was not due to ease of guessing the missing part of the sentence). The reversal of the bizarreness effect for cued as compared to free recall undercuts any simple account of the relationship between encoding time and recall. The advantage for bizarre sentences in free recall is thought to arise because distinctive associations between the bizarre content and experimental task provide a way of searching memory in response to a minimally informative retrieval cue. In contrast, the cued recall task eliminates the need for general search by providing informative cues about the material to be retrieved, but the usefulness of those cues depends on the creation of specific associations between the cue information and the information with which it appeared in the sentence. Cued recall is better for common than bizarre content because stronger memory representations are created when the action and patient make sense together. To test this idea, characteristic reading time patterns within the common and bizarre sentences were examined in relation to subjects' cued recall performance. For common sentences there was a positive correlation ($r = .40$, $p < .005$) between recall and second-pass reading on the modifier/patient, suggesting that a strategy that includes rereading this region strengthens its relationship to the instrument/action. In addition, overall memory was negatively related to the difference between bizarre and common sentences in second-pass reading after the patient was seen ($r = -.46$, $p < .001$). This relation indicates that a focus on rereading of the bizarre sentences detracts from creation of strong mental representations of the relations expressed in the sentences. How bizarre.

[1] Rayner, K., Warren, T., Juhasz, B. Liversdege, S. P. (2004). *JEP:LMC, 30*, 1290-1301.

***p < .001	Free Recall (n = 56)		Cued Recall (n = 56)	
	Bizarre	Common	Bizarre	Common
Whole-Sentence Reading Time (ms)	3966***	3452	3653***	3192
Sentence Access	31.7%***	13.0%	34.5%***	56.8%

The parallel computation of phrasal and nonphrasal constituents: Evidence from embedded adjectives in compound nouns

Cara Tsang, Craig G. Chambers (University of Toronto)
craig.chambers@utoronto.ca

It is well known that unfolding speech sounds are evaluated against stored lexical candidates and that incrementally-encountered words are rapidly integrated into an evolving compositional representation. But what happens when the input's relationship to these two processes is unclear? We explored this question in 3 visual world experiments using critical instructions such as *Click on the **blacksmith***. Of interest was the extent to which listeners interpreted the initial element of the target word as the beginning of the intended compound noun vs. an independent adjective that would plausibly form part of an emerging phrase (e.g., *the black$_{Adj}$ X$_N$*).

Properties of target and competitor objects were varied to explore the potential influence of pragmatic and semantic factors. In **Experiments 1 and 2**, the array of display objects on critical trials contained either a **contrastive competitor** (e.g., a black teapot as well as an orange teapot, where the use of an adjective is motivated), a **noncontrastive competitor** (e.g., a black teapot with no other teapot) or an **unrelated item** (e.g., a strawberry). Target images were manipulated across Experiments 1 and 2 such that they did/did not bear the color denoted by the embedded adjective (e.g., blacksmith wearing black vs. brown overalls), providing an additional means to diagnose an adjectival interpretation. Target names were recorded with typical "compound noun" prosody, potentially providing a biasing acoustic-level cue. Results from the two conditions with genuine competitors showed consideration of both target and competitor objects as the color portion of the compound was heard, suggesting that the two different linguistic solutions for the input were processed in parallel (and despite compound-biasing prosody). Importantly, fixations to the competitor were greatest in the contrastive competitor condition, indicating that a pragmatically-motivated modified phrasal description (e.g., *the black [teapot]*) was being considered and not simply that "black" was activated in the lexical cohort along with "blacksmith". However, the consideration of an adjectival interpretation was not fully dependent on the presence of referential contrast, as reflected in the significantly greater extent of fixations to a noncontrastive competitor compared to an unrelated item. An influence of the semantics of the embedded adjective was also apparent in target fixations, where for example a blacksmith with brown rather than black overalls had the effect of reducing early consideration of the target object. **Experiment 3** repeated the **unrelated item** and **noncontrastive competitor** conditions but in the latter case used competitor objects that would be strange to describe in terms of color due to its intrinsic/redundant nature for the object in question (e.g., a [black] tire). The color features of the target object were also manipulated (as was done previously across Experiments 1 & 2). Conditions with color-mismatching targets showed less competitor consideration compared to the earlier experiments, presumably reflecting a reduced expectation that color would be used in a phrasal description of the competitor object. However, competitor consideration was still greater than in the unrelated item condition, suggesting that an adjective analysis is still active in the early moments of processing even when the modifier would be both overdescriptive and redundant for the referent in question.

Together, the results highlight the fact that single word and phrasal solutions for temporarily indeterminate speech input are pursued in parallel. The results also reflect a range of interactions between bottom-up and top-down information sources. The influence of pragmatically-motivated contrast on form-class disambiguation appears to be weaker than what has been found in previous work using artificial lexicons (e.g., [1]), but the presence of pragmatic effects is nonetheless evident in the earliest moments, inconsistent with models that posit delays in the use of this information. We suggest the results are best understood as a finer-grained instance of "Merely Local" processing effects in which the local input is rapidly computed in multiple ways yet is nonetheless contoured by more global sources of information (e.g., [2]).

References

1. Magnuson, J., Tanenhaus, M.K.,& Aslin, R.N. (2008). *Cognition, 108*, 866-873.
2. Tabor, W., Galantucci, B., & Richardson, D. (2004). *Journal of Memory and Language, 50*, 355-370.

Individual differences in sentence processing: Separable effects of knowledge and processing skill
Peter C. Gordon, Wonil Choi, Renske S. Hoedemaker & Matthew W. Lowder (University of North Carolina at Chapel Hill)
pcg@email.unc.edu

Substantial effort has been devoted to understanding how variability in component cognitive abilities contributes to individual differences in reading skill. In sentence reading, variation across individuals is seen both in first-pass measures of early processing (e.g., skipping rates and gaze duration) and measures of later processing (e.g., regression-path duration). Recent research has attempted to understand how cognitive variation contributes to language comprehension by examining the relationship between tests of individual differences and a variety of reading time measures[1]. The current project pursues that strategy by examining eye-movement patterns in relation to two very different measures of individual differences: one that primarily taps knowledge and a second which primarily taps processing skill characterized as the ability to coordinate multiple cognitive processes in a sequential task.

The Author Recognition Test (ART) was used as a measure of knowledge obtained through print exposure[2]. In the ART, participants are given a list of authors' names mixed with foils and are instructed to indicate which of the names they recognize as being real authors. Processing skill was assessed using the Rapid Automatized Naming (RAN) task in which subjects are asked to name out loud 36 familiar objects arranged in a table as quickly and accurately as possible[3]; using four subtests (letters, digits, colors and objects). Data came from 340 undergraduate participants who participated in one of nine different eye tracking during sentence reading experiments in our lab. The correlation between ART score and average RAN completion times was very small ($r = -.13$) though this relationship was significant at the .05 level given the very large sample. The relationships between the individual differences and reading-time measures were assessed using mixed-effects models.

The figures below show reading measures across a range of word frequencies[4], dividing participants into three performance groups for ART and RAN (1: best, 3: worst). Variation in ART group was more strongly related to first-pass reading measures than was RAN group, as shown by the greater range and larger differences between groups on gaze duration (see figure), with similar results obtained for skipping rates. Additionally, very strong interactions between ART and word frequency were found for the first-pass measures. In contrast, variation in RAN group was more strongly related to regression-path duration than was ART group, as shown by the reversed pattern of range and group differences. These results suggest that variability in ART and RAN scores can be used to explain individual differences in two distinct components of sentence processing. Whereas increased print exposure seems to aid in the rapid identification of words, the ability to sustain attention during sequential processing, as measured by RAN, is associated with more consistent forward progress in reading.

References

[1] Kuperman & Van Dyke (2011), *Journal of Memory and Language.*
[2] Stanovich & West (1989), *Reading Research Quarterly.*
[3] Denckla & Rudel (1974), *Cortex.*
[4] SUBTLEXus, Brysbaert & New (2008), *Behavior Research Methods.*
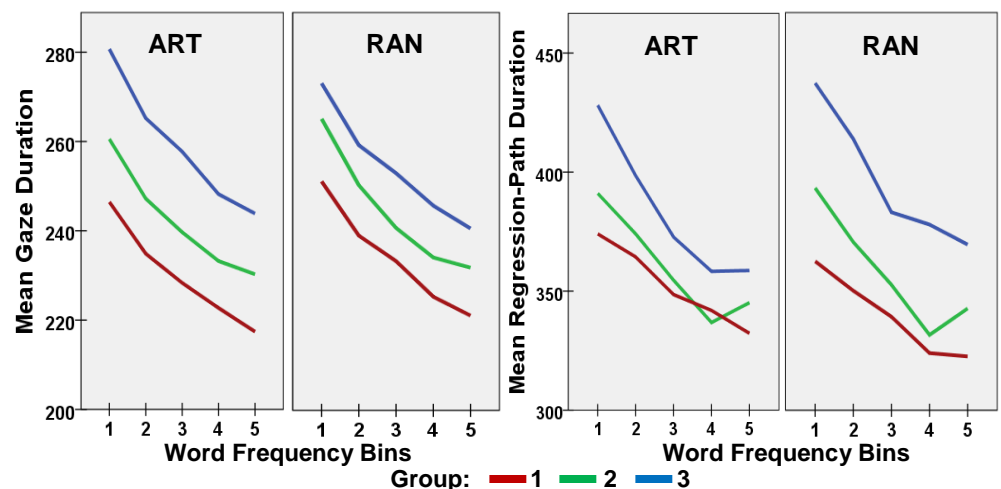
On the role of working memory capacity when prediction is not met: Evidence from NPI-processing
Juliane Domke (Humboldt University Berlin)
juliane.domke@staff.hu-berlin.de

There is growing consensus that readers/listeners use prior contextual information in order to build up a prediction of what word may come up next in context and that processing is costly when this prediction is not met. Moreover, recent findings suggest an influence of working memory capacity (WMC) when prediction is not met. That is, when WMC is low, higher processing demands are reflected due to the conflict between "keeping" prior build-up prediction and dealing with the mismatch due to actual input. [2] Using ERPs (event-related potentials), in our study, prediction was triggered on sentential context level using the German negative polarity item (NPI) *jemals* "ever" by comparing its appearance in an appropriate licensed with an ungrammatical unlicensed context (see Ex.1: *Kein/*Der Schüler hat* jemals *den Lehrer geärgert; lit. translated as „No/*The pupil has ever the teacher annoyed"*). The failure of NPI licensing has repeatedly yielded a biphasic N400-P600 processing pattern [3]. As to our study, we were interested whether to see an influence of WMC when the NPI is not licensed as prediction would not expect an NPI, when the appropriate licensor is absent. This is expected to reflect a load-related sustained negativity when WMC is low [i.e. 2, 4].

WMC [3] and EEG of 26 native German speakers were recorded. Participants were grouped according to their WMC test results in highWMC (n=14) and lowWMC (n=12) by median split. Figure 1 (below) shows an early negative going wave (300-500 ms) and a late positive going wave (600-900 ms) for both groups when the NPI is not licensed. Further, only low WMC group additionally yields a large late negativity (1100-1600 ms) as to the unlicensed condition. Results go in line with former findings on processing unlicensed NPIs showing a biphasic N400-P600 pattern. These correlates are fairly unaffected by WMC, as they occur in both groups. As was expected, influence of WMC is reflected by the large late negativity which only occurs in lowWMC group.

Results show that an unlicensed NPI triggers a mismatch with prior build-up prediction (in sentential context). This supports former findings in that the role of WMC influence does not affect anticipation of an unlicensed NPI but rather the processing of the mismatch. When prediction is not met, low WMC readers reflect a higher working memory load due to struggles of blocking prior build up prediction while processing the actual input.

Figure 1. Averaged ERP-waves of both WMC groups 0-1700 ms post stimulus (NPI *jemals*)
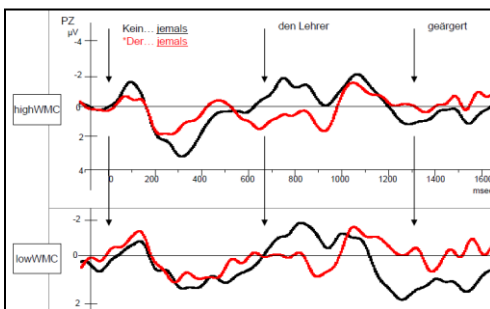
References. [1] Lewandowsky, S. et al. (2010) A working memory test battery for MatLab. *Behavior Research Methods, 42*, 571-585. [2] Otten, M. & J. van Berkum (2009), Does working memory capacity affect the ability to predict upcoming words in discourse? *Brain research*, *1291*, 92-101. [3] Saddy, D. et al. (2004), Processing polarity items: contrastive licensing costs. *Brain and Language, 90,* 495-502. [4] Vos, S. et al. (2001), Working memory constraints on syntactic processing an electrophysiological investigation. *Psychophysiology, 38,* 41-63.

Working memory and syntactic islands revisited

Edward Gibson (MIT), Greg Scontras (Harvard)
egibson@mit.edu

The unacceptability of syntactic islands like (1) and (2) (Ross, 1967; Chomsky, 1973, 1986) has been hypothesized to be due to factors unrelated to syntax, such as plausibility and working memory demands (Deane, 1991; Kluender, 1992; Sag et al., 2007; Hofmeister & Sag, 2010). In contrast, Sprouse, Wagers & Phillips (SWP, 2012) failed to observe correlations between working memory (WM) measures and acceptability ratings of syntactic islands like (1) and (2), and concluded that the unacceptability of island extractions must be due to grammar. However, three serious problems make SWP's results inconclusive: ***First***, very little data were collected for each participant (2 and 4 trials per condition in Experiments 1 and 2, respectively). This small number of trials plausibly led to poor estimates of an individual's rating of each condition, which could contribute to non-significant correlations. ***Second***, the control conditions for the island extractions were not matched to the island conditions for meaning, making the difference scores (used in the correlations) difficult to interpret. For example, the non-island long-distance extraction control for (1) is (3). But unlike (1), (3) is coherent in its meaning. A more appropriate, meaning-matched, control would include a long-distance dependency between the embedded position and a top-level position, but without wh-extraction, e.g., a long-distance pronominal dependency, as in (4). And ***third***, perhaps most importantly, SWP implicitly assume that processing costs are reflected in acceptability ratings, so that if an individual *X* experiences greater difficulty with some condition than individual *Y*, *X*'s acceptability ratings will be lower than *Y*'s. However, SWP do not show this to be true. Alternatively, consider an analogy between WM ability and physical ability. People of different physical abilities would judge a marathon to be more challenging than a walk in the park, in spite of their different abilities to accomplish these tasks. Similarly, an individual with low WM might be able to judge – as accurately as an individual with a high WM – which sentences sound easier vs. harder to process, even though low and high WM individuals may differ in their abilities to actually understand the sentences.

To address these limitations, we ran a study which included (i) a WM measure that has been shown to correlate strongly with fluid intelligence: complex sentence completion (Gibson & Fedorenko, 2012); (ii) a rating study of adjunct and NP islands, similar to SWP's, but with 8 trials per condition per participant, and with more appropriate control conditions, like (4) above. Furthermore, we included (iii) a rating study of nested vs. non-nested sentences, a contrast thought to be related to WM demands. In addition to ratings, participants answered comprehension questions about each sentence, providing a second dependent measure of sentence complexity. Two groups (each n=60) were run on Mechanical Turk, with similar results across the two groups.

Results. *First*, replicating previous results, we found stable within but varying across individuals completion rates. *Second*, this measure correlated reliably with the ability to answer comprehension questions (r=.495; p<.001). *Third*, as expected, robust differences in complexity ratings were observed for each contrast: nested vs. non-nested (2.12 vs. 3.17 out of 5; p<.001); adjunct-island vs. pronoun-dependency (1.72 vs. 2.22; p<.001); and NP-island vs. pronoun-dependency (1.89 vs. 2.27; p< .001). Critically, we found no correlation between our WM measure and the nested vs. non-nested difference score in the ratings (r =.052), suggesting that SWP's assumption about acceptability ratings reflecting WM demands is not warranted. It is thus unsurprising that SWP did not find correlations between acceptability ratings for island conditions and their WM measures. (We also see no such correlations.) Finally, the meaning-matched non-island control conditions were rated overall very low. We therefore propose that much of the difficulty in understanding island constructions has to do with people's ability to understand incoherent/implausibly backgrounded materials (cf. Erteschik-Shir, 1979; Goldberg, 2006, for similar arguments).

(1) *What do you sneeze if the dog owner leaves open at night?*
(2) *What did the chef hear the statement that Jeff baked?*
(3) *What do you hope that the dog owner will leave open at night?*
(4) *What is the thing, such that you sneeze if the dog owner leaves it open at night?*

How specific should I be? The optimal amount of information in online language comprehension
Si On Yoon, Sarah Brown-Schmidt (University of Illinois at Urbana-Champaign)
syoon10@illinois.edu

Speakers show sensitivity to the informativity of referential expressions given the local context (Brennan & Clark, 1996). However, informativity is not the only constraint on referring. Speakers regularly engage in informativity violations, including the use of previously entrained expressions that overspecify given the current context (Brennan & Clark, 1996). Less clear is whether *listeners* expect overinformativity. Some evidence suggests competition from previously-introduced subordinate labels (e.g., *carnation*) persists in subsequent contexts that allow basic terms (*flower*), based on increased looks at a carnation when interpreting "*Click on the car*" when "*carnation*" was used previously (Barr & Keysar, 2002). However, that effect might be due to experience clicking on the carnation, not due to overinformativity expectations. We examined whether listeners expect overinformative expressions when previously entrained in the discourse, while controlling for previous experience with the referent.

Experiment 1 examined whether addressees expect speakers to maintain collaboratively-established precedents despite the fact that the local context no longer warrants use of those expressions. Eye-tracked participants (P, *n*=32) followed an experimenter's (E) instructions to click on objects in a task with two phases: entrainment and test. In entrainment, P followed instructions to click on a target (e.g., striped shirt) and a competitor (e.g., checkered shirt) 4x each. Entrainment type (historical or ahistorical) was manipulated during entrainment trials. In the historical condition, E repeatedly used a modified expression to refer to both items (e.g., *Click on the striped/checkered shirt*), thus entraining specific terms. In the ahistorical condition, E referred to the items using locative phrases (e.g., *Click on the top left one*). Following entrainment, target displays contained only the target (the competitor was absent). We manipulated whether the target expression was overinformative (e.g., *Click on the striped shirt*) or appropriately informative (e.g., *the shirt*) for the current display. All manipulations were were within-subjects and the fillers contained multiple types of modified and non-modified noun phrases. If listeners expect entrained terms, on unmodified test trials, target fixations should be greater in the ahistorical condition, but on modified test trials, target fixations should be greater in the historical condition. *E1 Results:* Target fixations were numerically larger in the historical condition vs. the ahistorical condition for both unmodified expressions (*t*=1.013, *ns*), and for modified expressions (*t*=5.389, *p*<.01). The fact that the historical condition advantage was significant only for modified expressions is consistent with overinformativity expectations. However, the numerical advantage for the historical condition in the *unmodified* case suggests that the effect may be simply driven by familiarity with the critical noun phrase (e.g., [*striped*] *shirt*).

Experiment 2 used a cohort-competitor paradigm to test for expectations of listeners (*n*=32), controlling for familiarity. In entrainment trials, participants were exposed to pairs of exemplars from a target category (e.g., *dotted/argyle sock*) and a competitor category (e.g., *sitting/jumping dog*). In the historical condition, E used modified noun phrases, thus entraining modified terms (e.g., *dotted sock*). In the ahistorical condition, only locative phrases were used (e.g., *top left one*). Test trials included the target (dotted sock) and a cohort-competitor (jumping dog). On test trials, participants heard either modified (e.g., *dotted sock*—competes with "*dog*") or unmodified expressions (e.g., *dog*—competes with "*dotted*"). Both entrainment type (historical-ahistorical) and test type (modified-unmodified) were manipulated within-subjects. When interpreting unmodified terms (*dog*), the preference to fixate the target (dog) over the competitor (dotted sock) was higher in the ahistorical condition (*t*=1.869, *p*=.06), reflecting an expectation for a bare noun. Following modified expressions, target fixations were higher in the historical condition (*t*=-2.288, *p*<.05). This demonstrates that Ps expected entrained modified expressions only when they had been entrained in the discourse.

Conclusion. These findings demonstrate that, like speakers (Brennan & Calrk, 1996), *addressees* expect speakers to maintain the use of entrained precedents even when the local context changes and entrained terms are overinformative for the local context. These findings suggest that informativity (Grice, 1975) should be considered with respect to both the local and the historical context.

The realization of scalar inferences: Context sensitivity without processing cost

Stephen Politzer-Ahles (University of Kansas), Robert Fiorentino (University of Kansas)
sjpa@ku.edu

Scalar inference is the phenomenon whereby the use of a less informative term (e.g., *SOME*) is inferred to mean the negation of a more informative term (e.g., to mean *NOT ALL*). There is disagreement about how comprehenders arrive at the interpretation of *SOME* as *NOT ALL* during online comprehension. Default accounts assume that the inference-based *NOT ALL* interpretation is easy and automatic to realize, whereas context-driven accounts assume that it is effortful (see Katsos & Cummins, 2010, for a review). While numerous studies have investigated the speed of scalar inferencing, few have directly investigated the potential processing effort involved in realizing scalar inferences during online comprehension.

We adopt the self-paced reading paradigm of Breheny, Katsos, & Williams (2006) to test whether there are processing costs, in the form of reading time slowdowns, in sentences that are interpreted with a scalar inference versus those that are not. The results of that study are difficult to interpret because the stimuli differed in more ways than just the bias towards or against scalar inference—e.g., one condition involved a repeated noun penalty in the critical region (see Huang & Snedeker, 2009). We use materials (see (1)) that are maximally similar across conditions, differing only in the presence in the context sentence of the quantifier *all* or *any*. *All* creates an upper-bounded context that biases *some* to be interpreted as meaning *not all* (Katsos & Cummins, 2010), whereas *any* creates a lower-bounded context with no such bias. Control conditions involved the same manipulation but used *only some of* rather than *some of* as the quantifier, ensuring that the *not all* interpretation in these conditions is based on semantics rather than on pragmatic inference.

1) Mary was preparing to throw a party for John's relatives. She asked John whether (*all of them* / *any of them*) were staying in his apartment. John said that (<u>only some of them</u> / <u>some of them</u>) were. He added that <u>the rest</u> would be staying in a hotel.

Context-driven accounts predict slower reading times at or after "some of them" in the upper-bounded context (*all*) where the inference is realized, compared to the lower-bounded context (*any*). They also predict "the rest" to be read faster in the upper-bounded context, since *some of* in this context will be interpreted as *not all* and thus the reader will be aware of a remaining subset of referents, making "the rest" easier to integrate into the discourse model.

Consistent with the predictions of context-driven accounts, linear mixed models (N=27 participants) showed that "the rest" (appearing on average 1434 ms after *some of*) was read faster after the quantifier *some of* in upper-bounded contexts compared to lower-bounded contexts; there was no significant effect of context in this region or nearby regions for the control conditions. This suggests that the inference was ultimately realized in the upper-bounded contexts and not the lower-bounded contexts. Inconsistent with context-driven accounts, however, there was no effect of context in the reading times for "some of" or regions following it, providing no evidence that the inference was computationally costly. The numeric effect of context at the critical regions also did not significantly correlate with individual-level cognitive or pragmatic measures tested (cognitive control, working memory span, Autism Spectrum Quotient). These results suggest that scalar inferences are context-dependent but do not yield significant processing costs at the quantifier in this paradigm (for a constraint-based model that may account for these results, see Degen & Tanenhaus, 2011); these findings also motivate further tests for possible processing costs associated with inferencing using alternate methods.

References

Breheny, Katsos, & Williams, 2006. *Journal of Memory and Language, 100*, 434-463.
Degen & Tanenhaus, 2011. *Proc. 33rd Annual Conference of the Cognitive Science Society*, 3299-3304.
Huang & Snedeker, 2009. *Cognitive Psychology, 58*, 376-415.
Katsos & Cummins, 2010. *Language and Linguistics Compass, 4/5*, 282-295.

Eye movements reveal causes of delay in negative sentence processing

Ye Tian, Richard Breheny (University College London), & Heather Ferguson (University of Kent)
ye.tian.09@ucl.ac.uk

Many studies have established that negative sentences are more difficult to process than their positive counterparts. Here we present a visual world study exploring the reasons. Previous research suggest there may be at least two factors involved: accommodating context and inferring the actual state of affairs. [1] shows that shortly after reading a simple negative sentence, such as "the egg isn't cracked", comprehenders focus on a representation of the argument of negation, i.e. a cracked egg, before reaching a representation that is consistent with sentence meaning – a whole egg. This suggests that negative sentences are initially processed in the same way as their positive counterparts. However, [2] shows that representation of the positive counterpart is not mandatory. It is likely due to contextual accommodation. Negation is most frequently used for rejection or contradiction. When presented out of context, we probabilistically accommodate aspects of context for a negative sentence, including how the utterance is meant to be relevant. The most prominent context question or Question Under Discussion (QUD) for a simple negative is about the positive situation. However if a negative sentence has a negative prominent QUD (as in the negative cleft sentences used here), then participants do not first represent the affirmative counterpart.

The current study examines *when* contextual accommodation takes place during sentence processing, and at what point the meaning of the negation is incorporated. We studied this question using a visual world paradigm based on [3]. 33 participants heard simple negative sentences such as "John hasn't shut his dad's window" or cleft negative sentences such as "It is John who hasn't shut his dad's window", while looking at a scene consisting of a character in the centre and four objects: an open window (target), a closed window (competitor), and two distractors (an object in two states, such as a plain bagel and a bagel with cheese). The audio stimuli were recorded with natural prosody – stress on "hasn't" in simple negatives, and the name ("John") in clefts. There were 30 experimental sentences and 50 fillers. We counter-balanced sentence polarity and whether a sentence implies the initial or the end state of an action. One item was taken out in the analysis due to an error.

Figure 1 plots the log ratio of percentage of looks to the target/ competitor over time [4]. In the first 200ms window from the onset of the target noun (average noun duration: 629ms), bias to the target is significantly higher in the cleft than in the simple negative condition [$F1(1,32)=7.45$, $p=.01$; $F2(1,28)=10.78$, $p=.003$]. In the last 200ms window before the onset of the noun, there is a trend for the same effect (significant by items) [$F1(1,32)=2.41$, $p=.13$, $F2(1,28)=4.49$, $p=.04$]. This difference illustrates how different QUDs influence the processing of negative sentences. When the QUD is not in conflict with the sentence (as in the cleft case), participants are able to infer the true state of the object before the onset of the noun. Generally our results speak against the idea that the negation operator is only incorporated after the argument of negation is processed. In the simple negative case, there is no bias to either state, suggesting a conflict between the positive QUD and the inferred incomplete state.

References: [1] Kaup et al (2006). *QJEP*, *60*, 976–990. [2] Tian et al (2010). *QJEP,* 63, 2305–12. [3] Altmann & Kamide (2007). *JML*, *57*, 502–518. [4] Brown-Schmidt (2012) *Lang & Cog Proc*, *27*, 62–89.
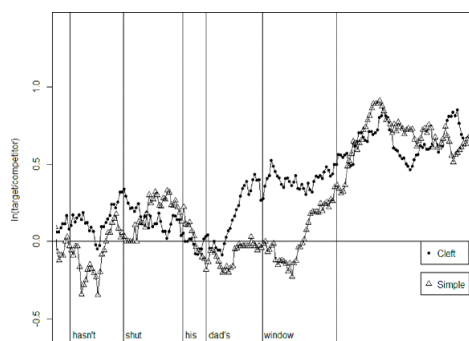
Figure 1: ratio of looks to target/competitor over time

Predictive computations underlie the N400's sensitivity to thematic role-reversals

Wing Yee Chow, Colin Phillips (University of Maryland) & Suiping Wang (South China Normal University)
wychow@umd.edu

Although previous studies have established that the N400 is larger for less expected words [1], many have observed that thematic role-reversals in S-O-V clauses (e.g., cop$_{SUBJ}$ thief$_{OBJ}$ arrest vs. thief$_{SUBJ}$ cop$_{OBJ}$ arrest) do not impact the N400 [2-4]. The N400's blindness to such anomalies has been taken to reflect a temporary 'semantic illusion' in which role-reversed sentences are initially perceived as plausible [2,3,5]. However, the N400's insensitivity might instead reflect that comprehenders' predictive computations have yet to be impacted by the structural role of the arguments. In line with this proposal, a recent study showed that the N400 became sensitive to role-reversal anomalies when the stimulus-onset asynchrony (SOA) between the object and the verb was increased from 600ms to 1800ms [6]. Here we tested two possible explanations for the N400's reemerged sensitivity. Did the N400 become sensitive to role-reversals because comprehenders had more time for predictive computations? Or did the extra time simply allow comprehenders to detect the implausibility at the verb more effectively [3]? We found that role-reversals elicited an N400 effect only when the verb was differentially predictable across conditions. We argue that the N400 is modulated by comprehenders' moment-to-moment expectations and that it becomes sensitive to role-reversals once argument role information is incorporated into comprehenders' predictive computations.

The present study (n=22) examined the role of predictive computations in the N400's sensitivity to thematic role-reversal anomalies. We used the S-O-V BA-construction in Mandarin Chinese and inserted a temporal phrase to increase the linear distance between the verb and its arguments (as in [6]). We created thematic role-reversals (implausible, 0% cloze) by reversing the structural roles of the pre-verbal arguments in canonical sentences in which the verb has high (62%) vs. low (3%) cloze probability, resulting in a 2 (high vs. low predictability) × 2 (canonical vs. role-reversed) design. If separating the arguments and the verb leads to N400 sensitivity because it gives the comprehender more time to detect role-reversal anomalies, then role-reversal anomalies should elicit an N400 effect regardless of the verb's cloze probability. Alternatively, if the N400's sensitivity critically depends on comprehenders' predictive computations, then role-reversal anomalies should elicit an N400 only when comprehenders expect the verb in the canonical control more strongly than in the role-reversed condition. Results revealed a significant N400 effect in the predictable condition, but the N400 remained blind to role-reversals in the non-predictable condition. In line with previous findings, role-reversals elicited a significant P600 in both conditions. Our results suggest that the N400 becomes sensitive to role-reversals when comprehenders have incorporated argument role information to anticipate the upcoming verb, whereas the P600 is more generally modulated by comprehenders' detection of an anomaly, and hence was elicited by implausibility in both role-reversed conditions. These findings show that the N400 provides an index of how predictive computations unfold during real-time comprehension.

References
[1] Kutas & Hillyard (1984). Nature. 307, 161-163.
[2] Kolk, Chwilla, van Herten, & Oor (2003). Brain & Lang. 85, 1-36.
[3] Hoeks, Stowe & Doedens (2004). Cog. Brain Res., 19, 59–73.
[4] Ye & Zhou (2008). Brain Res. 1203, 103-115.
[5] Kim & Osterhout (2005). JML. 52, 205–225.
[6] Chow, Phillips & Wang (2012). CUNY.

How hugging differs from giving a hug: Syntax, semantics or mapping
Eva Wittenberg & Jesse Snedeker (Harvard University)
eva@wjh.harvard.edu

Usually, semantic and syntactic arguments reliably align, allowing the parser to deduce semantics from the syntactic tree (Baker, 1988). One exception to this is the light verb construction, as in "Anne gave Julius a hug". ERP and behavioral studies have demonstrated that light verb constructions are processed differently than non-light constructions such as "Anne gave Julius a present" (Briem et al., 2010; Piñango, Mack & Jackendoff, forthcoming; Wittenberg & Piñango, 2011; Wittenberg et al., 2011). We investigate three hypotheses about the source of this processing difference:

A: The two constructions have different underlying syntactic configurations.

B: They have the same syntax, which leads to the unusual conceptualizations of 'give a hug' as a transfer event: [GIVE[ANNE,JULIUS,HUG]].

C: They have the same syntax, which in the case of a light verb construction maps onto conceptual structure in a noncanonical way, namely as a two-participant event ([HUG[ANNE,JULIUS]]) instead of a three-participant event ([GIVE[ANNE,JULIUS,HUG]]).

We test Hypothesis A using a standard syntactic priming task (Bock & Loebell, 1990), and B and C using a new task probing conceptual structure.

Experiment 1 tested Hypothesis A using a production priming paradigm. If light and non-light sentences have different syntactic structures, then priming across these sentence types should be smaller than priming within them. Participants read light or non-light ditransitive sentences (either double-object or prepositional-object) and then described a picture using either a light or non-light construction. Both light and non-light double-object sentences elicited more double-object picture descriptions than prepositional-object sentences ($z=-4.241$, $p<.01$), and did so to the same degree across and within construction type (no significant interactions). These findings show that the light and non-light datives share a common syntactic structure, eliminating Hypothesis A.

Experiment 2 used a sorting task to disentangle Hypotheses B and C. Hypothesis B predicts that light verb constructions are conceptualized as three role events (Source-Theme-Goal), while Hypothesis C predicts that they are conceptualized as two role events (Agent-Patient). Participants were trained to sort one-, two- or three-participant pictures according to number of roles. In the testing phase, participants were also confronted with written sentences. These could be light verb or base verb constructions, non-light transitive and ditransitive constructions (Julius loves Anne/Julius gives Anne a book), and intransitive or transitive joint actions (Julius and Anne meet/Julius meets Anne), with number of concrete and abstract objects varying. Joint actions were sorted consistently as two-roles, thus syntax alone was not a sorting strategy. Results show a three-way distinction between Agent-Patient events, events described by light-verb constructions, and Source-Theme-Goal events (all comparisons: $z<-.3.14$, $p<.01$). These results favor a middle ground between Hypotheses B and C, suggesting that the syntax in light verb constructions influences event representation. We discuss several alternatives for how this could take place.

Taken together, our results help explain reaction times and processing patterns in light and non-light constructions and make a contribution to our understanding of the relationship between syntactic and conceptual structure.

References. Baker, 1988. *Incorporation: A Theory of Grammatical Function Changing.* Bock & Loebell. 1990. *Cognition,* 35, 1-39. Briem et al. 2010. *Brain Research*, 1249, 173-180. Piñango, Mack, & Jackendoff. Forthcoming. *Proceedings of Berkeley Linguistics Society*. Wittenberg & Piñango. 2011.*The Mental Lexicon* 6:3, 393–413. Wittenberg et al. 2011. *CUNY Conference on Human Sentence Processing*.

Semantic effects on anaphor processing

Sara Peters[1,2], Timothy W. Boiteau[1] & Amit Almor[1] ([1]University of South Carolina, [2]Newberry College)
sara.ann.peters@gmail.com

Anaphor processing is affected by many factors including the anaphor's semantic relation to the antecedent. It is not clear whether and how semantic relations between other referents and the anaphor affect processing. In particular, it is unclear whether existing semantic relationships between the anaphor and other referents in the discourse affects the processing of noun and pronoun anaphors differently. In the current experiment, we asked whether mentioning two semantically related referents affects the processing of subsequent pronoun and repeated noun phrase anaphors to one of those referents.

Participants heard 3-sentence discourses describing the placement of referents within a visual display (see Table & Figure). Correct interpretation of each discourse was taken to mean that the grammatical subject of Sentence 1 (S1) was always considered to be the first referent spoken in Sentence 2 (S2) and Sentence 3 (S3) A verification task was used to identify participants' interpretation. S1 referred to a Target (*cat*) in relation to a Related (*mouse*) or Unrelated referent (*pump*). S2 referred to the Target in the subject position using a pronoun. S3 referred to the Target with a pronoun or repeated definite reference, then introduced a New referent (*bed*). This resulted in 4 conditions for experimental items (*n*=24), with participants (*n*=48) hearing one version of each amongst fillers constructed similarly, without the experimental manipulations (*n*=48).

Data were analyzed using Growth Curve Analyses to model participants' fixations to each item in the display as different conditions of S3 unfolded. The analyses indicated that following pronoun references Related referents were rejected faster than Unrelated referents. In contrast, following repeated definite references, Related referents were rejected more slowly than Unrelated referents. This was not observed for non-mentioned referents. Thus, previous mention and semantic relation affect the processing of pronoun and repeated definite reference differently.

These results show that when listeners encounter an anaphor, they activate the representation of the antecedent along with any referents previously mentioned with it. When one of those referents is semantically related to the antecedent, it is activated more strongly. Resolving the reference of repeated nouns is slowed when other referents are activated, possibly because they are considered as potential antecedents. Resolving the reference of pronouns is facilitated when other referents are activated, possibly because suppressing an alternative referent is easier the more strongly activated it is. Overall, these results support theories that highlight the importance of semantic memory representations in anaphor processing (e.g. Almor, 1999; Sanford & Garrod, 1981).

Sample Item	S1 Related Condition	S1 Unrelated Condition
S3 Pronoun Condition	S1. The cat is next to the mouse. S2. It is in the upper left corner. S3. It is above the bed.	S1. The cat is diagonal to the pump. S2. It is in the upper left corner. S3. It is above the bed.
S3 Repeated Condition	S1. The cat is diagonal to the pump. S2. It is in the upper left corner. S3. The cat is above the bed.	S1. The cat is diagonal to the pump. S2. It is in the upper left corner. S3. The cat is above the bed.
Verification Task	Verify (indicating True or False) that the sentences matched the Figure presented on the computer screen as the discourse unfolded. Sample Figure to the right.	

Target   Semantic Distractor

Sentence 1-Other   Sentence 3-mentioned
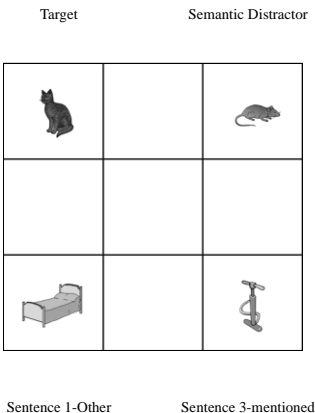
References

Almor, A. (1999). Noun-phrase anaphora and focus: The informational load
    hypothesis. *Psychological Review, 106*(4), 748-765.

Sanford, A. J. & Garrod, S.C. (1981). *Understanding written language: Explorations of comprehension beyond
    the sentence.* Chichester, England: John Wiley and Sons.

Advantages of extending vs. mixing metaphors: An ERP study

Les Sikos (Swarthmore College), Paul Thibodeau (Stanford University, Trinity College), Cassandra Strawser &
Frank H. Durgin (Swarthmore College)
lsikos1@swarthmore.edu

The widespread use of extended metaphors ("Crime has become an epidemic and there's no antidote strong enough to cure it!") suggests that it is advantageous to activate a mapping between two domains (CRIME, VIRUS) when interpreting metaphoric language. However, some have argued that spontaneous use of mixed metaphors ("Crime has become a beast that's roaring out of control and there's no antidote strong enough to cure it!") implies that such mappings may be suppressed for conventional metaphors (Glucksberg 2001). Here we used event-related potentials (ERPs) to test for online comprehension benefits of maintaining metaphoric consistency vs. potential costs of switching metaphors mid-stream.

Methods. Participants (N=96) read 108 stories containing two or more "background" sentences, followed by a target sentence (Table 1). 36 experimental stimulus sets were created such that background sentences described the same topic (e.g., crime) in one of three ways: (A) via a conventional conceptual metaphor (CRIME IS A VIRUS), (B) via an alternate conceptual metaphor (CRIME IS A BEAST), or (C) via a literal description. Target sentences containing readily interpretable extensions of A and B were created for each stimulus set. To assess semantic processing difficulty, we compared mean N400 (300-500 ms) and P600 (550-700 ms) amplitudes elicited by target words in three conditions: Match (AA, BB), Mix (AB, BA), and Literal (CA, CB). Participants saw 12 items per condition.

Results and Conclusions. ANOVAs showed that N400s were significantly more negative for Mixed and Literal conditions than for Match (ps<.05), with no differences between Mixed and Literal. These results suggest that the comprehension of extended metaphors was facilitated by first reading conceptually related conventional metaphors. Conversely, the same metaphorical extensions engendered greater semantic processing costs when preceded by either an alternate conventional metaphor or literal contexts. No reliable differences were found at the P600. However, ratings of individual items collected via Mechanical Turk (N=582) were used as predictors in a mixed-effects linear model. When activations at the N400 were taken into account, P600 amplitude was predicted by ratings of metaphoricity (p<.0001). These findings extend previous work suggesting that conceptual mappings in conventional metaphors are productive, and that this facilitation provides communicative efficiency specific to the conceptual mappings involved (Thibodeau & Durgin 2008).

Table 1. Example stimulus set
Background
   A. In big cities across America, crime has become an epidemic that can't be cured.
       It is beginning to infect small towns as well.
   B. In big cities across America, crime has become a beast that is roaring out of control.
       It is beginning to prey on small towns as well.
   C. In big cities across America, crime has become a problem that can't be solved.
       It is beginning to affect small towns as well.
Target Sentence (target word indicated by underscore)
   A. There is no <u>antidote</u> strong enough to cure it.
   B. There is no <u>cage</u> strong enough to restrain it.

References
Glucksberg, S. (2001). *Understanding Figurative Language: From Metaphors to Idioms*. Oxford University Press.
Thibodeau, P., & Durgin, F. H. (2008). Productive figurative communication: Conventional metaphors facilitate the comprehension of related novel metaphors. *JML*, 58, 521–540.

Lexically predicting visual features of word referents

Tristan Davenport, Seana Coulson, Vicky Tu & Benjamin Bergen (University of California, San Diego)
vwtu@ucsd.edu

Lexical prediction has recently become an important issue in sentence processing. However, it remains unclear what information contributes to lexical prediction, and at what level of specificity. Dikker and Pylkkänen (2012) have suggested in their Sensory Hypothesis that probabilistic expectations for orthographic word forms activate orthographic representations in visual cortex, facilitating expected words and interfering with unexpected words. However, comprehenders do not activate only orthographic information while processing language. Perceptual information about the objects and events being described is often activated during language comprehension (e.g., Zwaan et al., 2002). We report a behavioral experiment designed to link those two areas of research by testing whether sensory information activated during sentence reading can inform predictions of upcoming items. To that end, 50 participants read 108 sentences presented in rapid serial visual presentation (RSVP) format. All of the sentences chosen ended in words with a cloze probability above 80%, a type of context that known to encourage a predictive processing strategy (Federmeier et al., 2007). Half of the sentences were fillers ending in the expected word. In the other, critical items, the last word of the sentence was replaced with a picture, and the participant's task was to identify the picture aloud as quickly and accurately as possible. Critical stimuli consisted of 54 pairs of sentences, each ending in a picture, although participants only saw one of the pairs of sentences, depending on the experimental list. In each pair, the pictures corresponded to the same, predictable lexical item and also matched the unique visual features suggested by its corresponding sentence context. Critical sentence frames and pictures were combined so that each participant saw 18 items in each of 3 conditions: match, mismatch, and anomalous (see Table 1). Match pictures corresponded to both the predictable sentence-final word and to the visual features implied by the context. Mismatch pictures matched the predictable word but did *not* match the visual features implied by the context. Pictures in the anomalous condition did not fit the context at all. All pictures were counterbalanced across conditions in 6 stimulus lists, ensuring that across subjects, each picture appeared an equal number of times in each condition. Analyses of naming latency data revealed that participants named Match pictures 31 ms faster than Mismatch pictures and 167 ms faster than Anomalous pictures. Repeated measures ANOVA revealed significant differences among the three conditions ($F_1[2,98]=80.2$, $p<0.0001$; $F_2[2,178]=40.8$, $p<0.0001$). The followup analysis by subjects comparing only the Match and Mismatch conditions suggested that the Match pictures were named faster than the Mismatch pictures ($F_1[1,49]=7.2$, $p=0.01$; $F_2[1,89]=2.1$, $p=0.15$). Because subjects were significantly faster at naming a picture in the Match condition compared to the Mismatch condition in contexts that have been shown to facilitate predictive processing, we suggest that the participants completed the task by pre-activating likely visual features of the expected concept, in addition to amodal lexical information and orthographic word forms. These results extend Dikker and Pylkkänen's (2012) sensory hypothesis, suggesting that comprehenders not only make predictions of visual word-forms, but also of the visual features of expected word referents.

Table 1: A single sentence frame in the 3 critical conditions

Condition	Sentence Frame	Picture
Match	The ice hockey player put on his jersey and laced up his	
Mismatch	The ice hockey player put on his jersey and laced up his	
Anomalous	The ice hockey player put on his jersey and laced up his	

Semantic commitment in online verb processing
Nicholas Gaylord (The University of Texas at Austin), Micah Goldwater (Northwestern University),
Colin Bannard & Katrin Erk (The University of Texas at Austin)
nlgaylord@utexas.edu

It is well-known that word meanings change based on context, but there are still open questions regarding this process, regarding the timecourse of meaning resolution, the role of early commitments to interpretations while meaning-in-context is still being determined, and how these issues bear on the structure of mental semantic representations. We address these questions using a Speed-Accuracy Tradeoff (SAT) experiment on the resolution of intransitive verbs in the context of their subject nouns. McElree *et al.* (2006) report on an SAT study, finding that semantic composition in noun phrases is surprisingly slow, as evidenced by false positive responses to stimuli such as "Water pistols are dangerous" at delays of under 1 second. These findings also point to the presence of context-independent noun meanings that are activated prior to semantic integration. Here, we extend this approach to investigate the earliest stages of the semantic processing of verbs, motivated by findings that verbs are processed so rapidly as to potentially obviate the presence of a stored default meaning (Bicknell *et al.* 2010), as well as the fact that verb meanings have been shown to be more context-dependent than noun meanings (Gentner and France, 1988).

To begin addressing this question, we applied the SAT method to a verb meaning-in-context resolution task. Each verb appeared in the experiment with two subject nouns that cued different interpretations of that verb. Twenty-four participants were presented with stimuli containing a context sentence followed by a semantic probe (e.g. The dawn broke – Something shattered). On each trial, participants responded "true" or "false" after one of 8 delays between 200 to 3000 ms following presentation of the probe. We hypothesized that, while verb-argument semantic composition might be more rapid than noun-modifier composition, verbs would nonetheless carry default interpretations and that this would be evidenced by reduced response accuracy at short delays to stimuli such as the example just given, relative to response accuracy for semantically impossible controls.

We analyzed our data using mixed effects logistic models (including participant and verb as random effects on all nested items) with response as an outcome, and an interaction between correct response and response delay as predictors. Response accuracy was significantly lower at shorter delays for false-given-the-context probes, but participants remained accurate in rejecting false-regardless-of-context controls. We take this processing difference as an indication that verbs do in fact have default meanings activated prior to composition.

The effect of this default meaning is further evidenced by a more detailed treatment of sense dominance: the highest error rate was observed at short delays when responding to stimuli in which the context sentence activated a subordinate sense of the verb, while the probe corresponded to that verb's dominant sense. While the strength of a verb's default meaning might be expected to depend on verb imageability, we found no such effect but did find an effect of the distributional association between the verb and the semantic probe. This suggests that verbs' default meanings may be better characterized via distributional association rather than their imageable properties.

References
Bicknell, K., Elman, J., Hare, M., McRae, K., & Kutas, M. (2010). Effects of event knowledge in processing verbal arguments. *Journal of Memory and Language* 63, pp. 409-505.

McElree, B., Murphy, G., & Ochoa, T. (2006). Time course of retrieving conceptual information: A speed-accuracy tradeoff study. *Psychonomic Bulletin and Review* 13, pp. 848-853.

Gentner, D. & France, I.M. (1988) The verb mutability effect: Studies in the combinatorial semantics of nouns and verbs. In Small, S.L. *et al.* (eds.) *Lexical Amgibuity Resolution*, Kaufman, pp. 343-382.

Regeneration in verb phrase ellipsis resolution
Suzanne Belanger & Ron Smyth (University of Toronto)
suzanne.belanger@utoronto.ca

We report on experimental findings investigating the resolution of verb phrase ellipsis (VPE), which a number of researchers have argued involves accessing a 'surface representation' of the antecedent (e.g., Tanenhaus et al. 1985; Mauner et al. 1995, among others). However, Potter & Lombardi (1990,1998) argue that fully formed 'surface' representations are not actually retained in working memory, and that verbatim recall involves accessing a conceptual representation and regenerating the to-be-recalled (TBR) sentence using recently activated lexical items and syntactic frames. The main piece of evidence for this is 'lure intrusion': when presented with a synonym in a list after the TBR sentence, subjects sometimes replace the original word with the synonym, even if this requires syntactic changes. We propose that resolution of VPE is essentially like ('silent') recall, accessing a conceptual representation and then regenerating the missing VP. To test this, we presented 112 participants with materials of the following type for self-paced reading:

John gave the money to the church/gave the church the money (NP-PP or NP-NP)
##donated/kissed (Lure or no lure) (MASKED PRIME)
His daughter was angry that he did. (ELLIPSIS SENTENCE)
John gave the money to the church/the church the money. (VS- Match or Mismatch)

The antecedent contains a dative verb that can take either an NP-PP or NP-NP complement. The masked prime is either a lure--a verb synonymous with the original verb but restricted to only taking NP-PP -- OR an unrelated verb. Then follows the ellipsis sentence, and a Verification Sentence (VS) which either matches or mismatches the antecedent. Previous studies manipulating particle verb word order (Tanenhaus et al. 1985) showed an effect of Match on VS RTs, with slower RTs associated with mismatching VSs following an ellipsis. We expected to see a replication of this Match effect in all but one comparison: NP-NP antecedent with Lure. If the lure introduces the NP-PP structure, then this will sometimes be used in resolving the ellipsis. If so, then this would not clash with an NP-PP verification sentence, despite the antecedent having an NP-NP structure.

This is exactly the pattern that was found - we found a significant 3-way interaction for PP/NP x Lure/no lure x Match/MM (p<.04). We divided the data and examined Lure and No Lure separately. The No Lure data showed only an effect of Match, as we expected based on previous findings. However, the Lure data showed a significant interaction of PP/NP x Match/MM (p=.002). Further analysis revealed a significant Match effect for NP-PP but not for NP-NP. If in resolving the VPE, participants accessed a verbatim representation of the antecedent, the usual Match effect would have been observed regardless of the presence of the lure. Instead, these results are consistent with the hypothesis that resolution of VPE, like recall, starts at the conceptual level and regenerates using recently activated lexical items and syntactic frames – the latter not always the one introduced by the antecedent. We are currently running a second experiment using the same materials but replacing VPE with the so-called 'deep' verbal anaphor "do it" for comparison; preliminary results show no effect of lure on RTs, supporting the idea that the presence of the ellipsis is critical in giving rise to the observed pattern.

References. [1] Mauner, G., M. Tanenhaus & G. Carlson (1995) 'A note on parallelism effects in processing deep and surface verb-phrase anaphora' Language and Cognitive Processes,10:1,1-12. [2] Potter, M. & L. Lombardi (1990) 'Regeneration in the short-term recall of sentences.' Journal of Memory & Language, 29, 633-654. [3] Potter, M. & L. Lombardi (1998) 'Syntactic priming in the immediate recall of sentences' Journal of memory & language, 38, 265-282. [4] Tanenhaus, M., G. Carlson & M. Seidenberg (1985) 'Do listeners compute linguistic representations?' in D.R. Dowty, L. Karttunen and A.M. Zwicky (eds.), Natural language parsing: theoretical, psychological and computational perspectives. Cambridge: Cambridge University Press

Two flavors of long distance dependency discerned through island effects

Dan Parker & Bradley Larson (University of Maryland)
dparker3@umd.edu

Research on the processing of filler-gap dependencies has shown that the parser effectively deploys island constraints according to the grammatical details of the constructions used [1,2]. We present evidence from a speeded acceptability judgment study which suggests that speakers respect island constraints in the first conjunct of conjoined sentences with across-the- board (ATB) wh-extraction (1), but not the second. These results lead to the conclusion that the two conjuncts form different types of dependencies with the wh-phrase, once grammatical properties particular to conjoined sentences are taken into account. ATB constructions contain a single wh-phrase that enters into multiple dependencies, as in (1). While previous accounts assume that ATB constructions involve syntactically parallel gaps [3,4], such accounts cannot explain certain asymmetries, e.g., anaphor reconstruction is permitted in the first conjunct, but not the second, as in (2). We hypothesize that this asymmetry arises because the second conjunct does not form a syntactic relationship with the wh-phrase, and thus does not trigger syntactic reconstruction. Instead, the second conjunct forms a purely semantic relationship with a syntactically-unrealized variable associated with the wh-phrase. This hypothesis predicts that the parser should posit a gap in the second conjunct of an ATB construction even if that gap occurs inside an island. The processing difficulty associated with the perception of a syntactic island violation should be reduced or eliminated for the non- syntactic dependency between the gap in the second conjunct and the wh-phrase. In a speeded acceptability judgment task, we manipulated ISLAND POSITION (first vs. second conjunct) in conjoined sentences with ATB wh-extraction (sample sentences: a-b), and ISLAND PRESENCE (present vs. absent) in sentences with a single filler-gap dependency (c-d) to provide a baseline for island sensitivity. Sentences were presented word-by-word with a 400 ms SOA, after which participants (n=17) gave a binary acceptability judgment. Logistic mixed-effects modeling revealed a significant effect of ISLAND PRESENCE for sentences with a single filler-gap dependency: sentences with extraction out of an island were unsurprisingly less acceptable than sentences with licit extraction (p=0.01). Within ATB sentences, modeling showed a significant effect of ISLAND POSITION: sentences with an island in the second conjunct were more acceptable than sentences with an island in the first conjunct (p=0.04). The same contrasts were found in an untimed, offline rating study (n=18). These results support our hypothesis that the second conjunct does not form a syntactic relationship with the wh-phrase, but rather a semantic relationship. They may also militate against the view that island constraints are epiphenomena of processing factors related to the complexity of the filler phrase [5]. Filler phrase complexity is held constant in the present study, yet differential processing costs are found. Taken together, these results support recent findings that the parser deploys detailed grammatical constraints particular to coordinated phrases to guide decisions about where it may complete a filler-gap dependency [2].

(1) What did Sam bake and Tom eat? (2) *[Which picture of herself] did Tom sell and Sue buy?

	Sample sentences	z-score	% accepted
(a)	Who was it that Julie wondered whether William admired and Vicky knew that Iris respected?	0.25	71%
(b)	Who was it that Vicky knew that Iris respected and Julie wondered whether William admired?	0.53	85%
(c)	Who was it that Julie wondered whether William admired?	0.39	78%
(d)	Who was it that Vicky knew that Iris respected?	0.78	98%

References. [1] Phillips. 2006. *Lang.* [2] Wagers & Phillips. 2009. *JL.* [3] Citko. 2005. *LI.* [4] Nunes. 2004. Linearization of Chains and Sideward Movement. [5] Hofmeister & Sag. 2010. *Lang.*

Feedback, risk sensitivity and response-contingent financial payoffs affect reading time for syntactically ambiguous sentences

Luis Chacartegui-Quetglas & Colin Bannard (University of Texas)
luischq@utexas.edu

Swets, Desmet, Clifton, & Ferreira (2008) found that the time taken to read ambiguous regions of sentences is task-dependent, with reading time being longer when disambiguation might aid performance on a subsequent task. We here explore whether such task dependence might be modeled using economic concepts that have proved valuable in other domains of cognitive psychology.

Given a choice between options, people combine the value of each outcome with the probability of its occurring (albeit in suboptimal ways). Furthermore, payoffs and outcome probabilities have been shown to affect time spent in information seeking (Diederich & Busemeyer, 2006). We conceptualize reading as information seeking, and explore the impact of potential economic losses and of outcome (interpretation) probabilities on the reading time for ambiguous sentences. We created 24 sets of three relative-clause-containing sentences that differed in ambiguity – e.g. *the [classmate/chihuahua/bike] of the boy that plays catch is in the garden.* 50 participants rated the probability of each sentence interpretation on a 1-7 scale. We used these ratings to assign a probability to each interpretation of each sentence using Luce's choice axiom.

In a first study, 24 participants read one sentence from each set while their gaze was monitored using a Tobii X120 eye tracker. After each sentence they were asked a disambiguation question (e.g. *who played catch?*) and chose between the two attachment-candidate NP's. 20 points were given for a correct answer. The penalty for an incorrect answer varied in blocks (1,10, 50 or 100 points). The participant with the highest score received a gift certificate to spend at the University store. The order of the blocks, and their combination with sentences was counterbalanced for each participant. English speakers are known to have an NP2-attachment bias for relative clauses. We found that total reading time for the disambiguating region (e.g. *that plays catch*) increased as the probability of NP1-attachment (the non-preferred interpretation in the absence of other information) increased, and that this effect was stronger with higher points at stake (and thus the expected loss if the participant were to give the default NP2-attachment response was greater). We also obtained risk aversion (Arrow Pratt) ratings from each participant using a simple survey. We found that participants' risk aversion modulated the effect; reading time increased more for more risk-averse people.

A second set of 36 participants read the same sentences and questions. For this version of the study, participants received feedback about their response for the current trial as well as a running total. 1 point was given per correct answer and the penalties for an incorrect answer were 1, 2, 4 and 8. Participants started with 0 points which would earn them $2.5 at the end of the session. A positive score earned them up to $5 and a negative score from $2.5 to $0. The addition of feedback radically affected reading strategies. Overall reading times for the ambiguous region were significantly longer. Entropy (highest where NP1 and NP2 probabilities are equal) was found to be a better predictor of reading time than NP1 probability, suggesting that readers adapted to the 50% rate of NP1-correct items revealed by feedback, overriding their NP2 bias. The effect was stronger for trials that followed the participants being informed of an incorrect response and their associated loss. The magnitude of the loss did not have an effect, which we attribute to participants already being at ceiling following error. Consistent with the findings from the first study, reading time increased more as a function of entropy for those participants with high survey-revealed risk aversion.

Refereneces
Diederich, A. & Busemeyer, J. R. (2006). Modeling the effects of payoff on response bias in a perceptual discrimination task: Bound- change, drift-rate-change, or two-stage-processing hypothesis. *Perception & Psychophysics*, 68, 194-207.
Swets, B., Desmet, T., Clifton, C. Jr. & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence self-paced reading. Memory & Cognition, 36 (1), 201-216.

A rational account of regressions in syntactically complex sentences

Klinton Bicknell & Roger Levy (UC San Diego)
kbicknell@ucsd.edu

Why do readers often make regressions to previous parts of a sentence when faced with syntactic difficulty, such as in garden path sentences? The dominant answer to this question is that making a regression helps with syntactic reanalysis (e.g., Frazier & Rayner, 1982). Levy (2008) suggested an alternative explanation. Readers may maintain uncertainty about the identities of previous words, and update their beliefs about these words on the basis of new input. Under this hypothesis, a word that is unlikely in its syntactic context may cause a large update in beliefs about previous words. Levy suggested that readers use a heuristic of making a regression when this update (termed the error identification signal, EIS) is large. In support of this account, Levy et al. (2009) reported an experiment with sentences as (1)–(4). The difficulty contrast between sentences 1–2 at the verb tossed/thrown was previously argued by Tabor et al. (2004) to reflect syntactic difficulty, but under Levy's account, is because 'tossed' induced a large EIS relative to 'thrown', specifically, suggesting that 'at' should have been 'and' or 'as'. Supporting this view, Levy et al. found that the difference in regression rates between (1)–(2) was markedly reduced between (3)–(4), which did not contain 'at'.

Bicknell and Levy (2010) presented a rational model of eye movement control in reading, in which a reader uses Bayesian inference to combine noisy visual input with a probabilistic language model to infer a sentence's form and structure. They used this model to derive a new, general explanation for regressions in reading, closely related to the EIS heuristic. Specifically, they presented simulations showing that making occasional regressions to previous words when confidence in their identity becomes low allows a reader to read faster and more accurately. Bicknell and Levy, however, performed these simulations on syntactically simple sentences, and thus, it is unclear whether this account can explain regressions elicited by syntactic difficulty.

The present work addresses this question by extending Bicknell and Levy's model to use a syntactic language model: a small probabilistic context-free grammar with rule probabilities estimated from the parsed Brown corpus. Simulation results (top Figure, with Monte-Carlo-derived error bars) showed that the model made regressions that qualitatively mirror those in Levy et al.'s human data (bottom Figure). The rate of first-pass regressions from the critical verb showed a significant interactive pattern, with the most regressions in condition (1). The same pattern was seen in the rate that a regression from the verb eventually resulted in a fixation on the preposition ('go-past regressions'), though only marginally significant.

These results demonstrate that Bicknell and Levy's account of regressions can explain (at least one class of) long-range regressions elicited syntactically. In addition, since Bicknell and Levy showed that these regressions increase reading efficiency for syntactically simple sentences, the result raises the prospect of a unified explanation of long-range and short-range regressions, and broadens the range of eye-movement phenomena in reading that can be understood as adaptive action in the face of uncertain input.

1) The coach smiled at the player tossed a Frisbee.
2) The coach smiled at the player thrown a Frisbee.
3) The coach smiled toward the player tossed a Frisbee.
4) The coach smiled toward the player thrown a Frisbee.
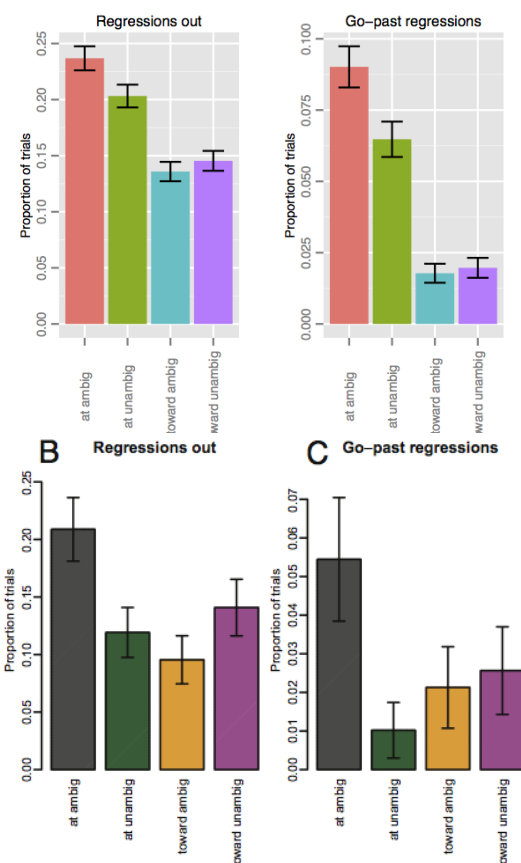
Online filler-gap dependency formation and that-trace effect
Morgan Purrier, Rebekah Ward, Lauren Ackerman & Masaya Yoshida (Northwestern University)
MorganPurrier2013@u.northwestern.edu

Introduction. The processing of the wh-filler-gap (WhFG) dependency formation has a long history in the sentence processing literature. It has been established that the parser does not try to posit a gap in certain domains that are not grammatically sanctioned, e.g., islands([1,2]). This study aims to show that the online WhFG dependency formation process is sensitive to other types of constraints as well. We show that the parser is sensitive to the so-called that-trace filter ([3]), for the wh-movement from right after the complementizer *that* is illicit in English. Though, the nature of the that-t filter itself has been controversial ([4]), it is clear that the presence of that is a crucial factor in the contrast in (1).
(1) Fred forgot who Aisha said (*that) t_{who} would write to when he got a job.

Experiment. Assuming an active gap search for the wh-phrase ([1]) and this search process is grammatically constrained, then we expect that the parser does not try to posit a gap right after *that*. We employed the Filled-Gap Effect (FGE) paradigm ([1]) in word-by-word moving window experiment by manipulating the following two independent factors in a 2×2 factorial design: Wh Type (Wh-NP vs. Wh-PP)×Complementizer (that vs. null). The sample set of experimental stimuli is summarized in **Error! Reference source not found.**).

(2) a./b. Fred forgot [who] Aisha said **that/ø** *he* would write to __when he got a job.
(2) c./d. Fred forgot [to whom] Aisha said **that/ø** *he* would write   __when he got a job.

If the parser tries to terminate the WhFG dependency, as soon as possible, while respecting the that-t filter, then we would expect that the parser would try to posit the gap in the embedded subject position when *that* is absent, as this position is the closest potential gap position. However, if *that* is present, then the parser avoids positing a gap after *that* because it leads to a that-t filter violation. We would expect the reading time slowdown associated with the FGE at the embedded subject position, *he,* in the wh-NP/ø condition in 0b) but not in the wh-NP/that condition 0a). In the Wh-PP condition in 0c/d), we do not expect that the parser tries to posit a gap for the Wh-PPs in the embedded subject position because a Wh-PP like, *to whom,* is never a subject in English, therefore no FGE, in comparison. We checked the acceptability of all the sentences in order to check whether or not these sentences are globally grammatical. A 7-point-scale acceptability-rating task revealed that all the stimuli show equal acceptability (Fs<1). We also conducted a word-by-word self-paced moving window experiment (n=40). There was a significant interaction of Wh-type x Complementizer (ps<.05) at the embedded auxiliary region (*would* in 0)), which is one region after the embedded subject pronoun. The pairwise comparison revealed that the embedded auxiliary was read significantly slower in 0b) compared to 0a) (ps<.05), but no such difference was detected in Wh-PP conditions. We found no significant effects in other regions. We take the reading time difference detected at the embedded Auxiliary as the effect carried over from the embedded subject region, as there are no other differences among the conditions other than the manipulated independent factors.

Discussion. The results suggest that the parser did not try to posit a gap right after the complementizer *that*, signifying that, like other island constraints, the parser respects a that-t filter on WhFG dependency formation. One may argue that the presence of the complementizer *that* can signal the presence of the subject, and the reading time slowdown is due to the parser not expecting the subject in the embedded clause when *that* is absent. Such interpretation, however, is not tenable. If the "unexpected" subject is the cause of the slowdown, we should expect the same in the Wh-PP conditions. However, there was no main effect in terms of the Complementizer factor. Thus, the unexpected subject cannot be the cause of the slowdown in the Wh-NP/ø condition. This result further supports the argument that Wh-phrase is **not** directly associated with the subcategorizers ([5,6]), as the subject position comes before the verb/preposition is processed.

References. [1] Stowe (1986) *LCP* 3:227-245. [2] Phillips (2006), *Language* 82:795-823. [3] Chomsky & Lasnik (1977), *LI* 8: 425-504. [4] Featherston (2004) Lingua *115*, 1277-1302. [5] Lee (2004) *JPR* 33:51-73. [6] Aoshima et al., (2004) *JML* 51:23-54.

The role of morphology in phoneme prediction: Evidence from MEG

Allyson Ettinger, Tal Linzen & Alec Marantz (New York University)
ake1@nyu.edu

Reaction times in behavioral auditory word recognition show effects not only of the point at which a morphologically complex word becomes uniquely distinguishable from all other words in the language, whether morphologically related or unrelated ("kindly" from "kite" and "kindness"), but also of the point at which its stem ("kind-") becomes distinguishable from other stems ("kindly" from "kite") [1]. These results weigh against full-form models of word processing in favor of models assuming a role for obligatory morphological decomposition.

Balling and Baayen suggest that these effects may be attributable in part to the low predictability of the phonemes at these uniqueness points. Indeed, the sentence processing literature suggests that less predictable words, which convey more information (carry high surprisal, [2]), are read more slowly [3]. Recent neural evidence suggests that auditory cortex activation at a phoneme is likewise correlated with the surprisal of that phoneme [4].

The present study leverages the temporal resolution of magnetoencephalography (MEG) to probe the role of morphology in phoneme prediction in the auditory cortex. Thirteen participants performed a lexical decision task on stimuli consisting of a) bimorphemic words paired according to initial morpheme ("clockwise", "clockwork"), and b) monomorphemic words paired according to shared phonological material starting at onset ("bourbon", "burble"). Surprisal values for each phonological segment of each stimulus were estimated based on corpus frequencies, and the Penn Forced Aligner was used to time-lock these surprisal values to the segment boundaries within the recorded stimuli.

Correlational analysis of activation in the left transverse temporal gyrus showed a highly significant effect of phoneme surprisal, such that less predictable phonemes were associated with significantly higher processing cost ($p < .001$). This effect was most pronounced at the end of the word. Importantly, the effect of surprisal was significantly stronger in bimorphemic than in monomorphemic words ($p < .05$). This difference between morphological conditions indicates that morphological information plays a role in generating phoneme predictions over and above the role attributable to the distributional properties of the phonemes alone. Our results furthermore lend support for the role of higher levels of linguistic knowledge in generating predictions about upcoming sensory input [5].

Figure 1. Timepoint by timepoint correlation between phoneme surprisal and auditory cortex activity derived from linear mixed effects model (contrasting bimorphemic and monomorphemic conditions).
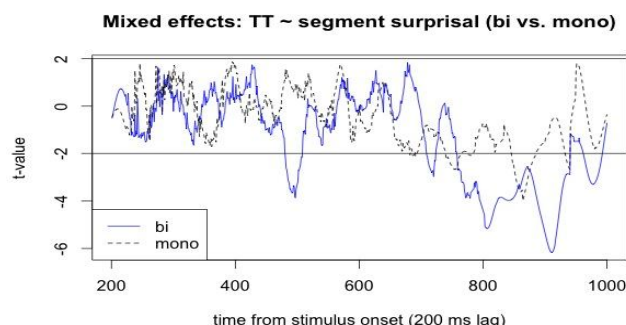
References
[1] Balling, L., Baayen, R.H. (2012), *Cognition 125*, 80-106. [2] Hale, J. (2001), *Proceedings of NAACL 2*, 159-166. [3] Levy, R. (2008), *Cognition 106*, 1126-1177. [4] Gagnepain, P. et al. (2012), *Current Biology 22*, 615-621. [5] Dikker, S., et al. (2009), *Cognition 110*, 293-321.

Verb-argument processing with and without event-related knowledge impairment
Michael Walsh Dickey (University of Pittsburgh, VA Pittsburgh Healthcare System) & Tessa Warren (University of Pittsburgh)
mdickey@pitt.edu

Event-related conceptual knowledge rapidly affects verb-argument processing (Kuperberg, 2007; McRae, et al. 2005). Verbs activate event representations in semantic memory, which in turn activate common event participants (McRae, et al., 2005). This knowledge allows comprehenders to predict likely upcoming arguments of verbs (Metusalem, et al, 2012), and creates rapid disruption when the arguments violate these expectations (Rayner, et al., 2004). However, it remains controversial whether activation of event-related knowledge is required for verb-argument processing, or whether some aspects of verb-argument processing may be independent of such knowledge (Boland, 2005).

This study provides novel evidence from aphasia demonstrating the limits of event-related conceptual knowledge in verb-argument processing. Adults with left-hemisphere damage and aphasia (n=8) read stimuli from Warren and McConnell (2007) in two self-paced reading studies. Sentences contained a critical NP (underlined), which was either a plausible argument of the verb [1a], a highly implausible/unlikely but possible argument [1b], or an impossible argument that violated the verb's selectional restrictions [1c]. In Study 1, participants provided sentence-final acceptability judgments. In Study 2, they answered sentence-final comprehension questions. Half the participants had a conceptual-semantic deficit, with impaired event-related representations as indicated by poor performance on both a standardized measure of conceptual semantic processing (Pyramids and Palm Trees; Howard & Patterson, 1992) and a novel test of event-related conceptual knowledge, in which they had to judge whether depicted events "might normally happen" (pictures from Proverbio, et al, 2009). The other participants did not have such conceptual-semantic deficits.

In both Study 1 and 2, the participants <u>without</u> conceptual-semantic deficits were particularly sensitive to event-related information. Their reading times did not differ across conditions at the critical noun (carrots) but sentence finally, the implausible condition [1b] was slower than both the plausible [1a] and impossible [1c] conditions. Acceptability judgments were most accurate in the implausible condition, but comprehension question accuracy was reduced for both implausible and impossible conditions. This pattern suggests that violations of event-related expectations more strongly affected performance than selectional-restriction violations for this group.

Strikingly, participants with conceptual-semantic impairments showed relatively intact rapid processing of language-related information (including verb-specific knowledge like selectional restrictions). Their reading times were slower in both the implausible [1b] and impossible [1c] conditions at the critical noun, with more disruption for the impossible condition in Study 2. Interestingly, this group's performance on acceptability judgments (Study 1) and comprehension questions (Study 2) was better and did not differ across conditions. This pattern suggests that event-related knowledge may not interfere with the mental models these adults build

The results from participants with semantic impairments suggest that verb-argument processing can proceed normally even when event-related conceptual knowledge is impaired. This is surprising if verb-argument processing always requires (or may be reduced to) activation of event-related conceptual knowledge (Matusalem, et al., 2012). However, findings from semantically-unimpaired participants suggest that event knowledge may be critical to verb-argument processing when processing of other verb-argument information (such as selectional restrictions) is impaired.and between bullet points, single-spaced is preferred.

(1) Stimuli from Warren & McConnell (2007; | marks self-paced presentation regions)
a. Maria | used | a knife | to chop | the | large | carrots | before dinner | last night.
b. Maria | used | some bleach | to clean | the | large | carrots | before dinner | last night.
c. Maria | used | a pump | to inflate | the | large | carrots | before dinner | last night.

MEG evidence for immediate reference resolution within a visual world

Christian Brodbeck & Liina Pylkkänen (New York University)
christianbrodbeck@nyu.edu

An elementary operation in language comprehension is identifying the entities that individual expressions refer to. Eye tracking studies suggest that people find the referents of noun phrases (NPs) quickly and efficiently [1]. Less is known about the neural implementation of this process. We report an investigation of the neural processing of simple, referentially unambiguous adjective-noun phrases in a magnetoencephalography (MEG) study inspired by the visual world paradigm.

Our study builds on three prior results. First, in the absence of a referential task, the combination of adjectives and nouns has been shown to involve the left anterior temporal lobe (ATL) and ventromedial prefrontal cortex (VMPFC) [2]. Second, EEG studies have shown that referentially ambiguous expressions elicit an enhanced frontal negativity, suggesting increased processing cost. Third, an fMRI study found that referentially ambiguous NPs activated a network of brain regions including VMPFC, but not ATL [3]. Together these results might suggest that the left ATL is involved in combinatory language processing, and the VMPFC is involved in how these processes feed into the process of reference resolution.

In order to probe the time course of successful reference resolution in simple expressions, we adapted a visual world paradigm for MEG. In each trial, participants (N = 16) first saw a constellation of three simple figures, which was then replaced by a question presented word for word (*Was the green circle next to the triangle?*). In the questions, shapes were referred to with color-shape NPs. The target expression was the first adjective of each question. In half the questions, there were two items matching the color denoted by the adjective; in the other half, the adjective picked out a unique entity (i.e., there had been only one item with this color). Based on earlier eye tracking studies [1] we expected participants to resolve reference incrementally: if the adjective only fit one item, participants should resolve reference soon after reading the adjective; if the adjective was compatible with two items, they should have to wait for the noun.

We extracted epochs related to the onset of the relevant adjective, and transformed them to source space using MNE. We extracted the time course of activation for specific regions of interest in the ATL and ventral prefrontal cortex and tested for temporally contiguous clusters where the two conditions differed using permutation cluster tests [4]. Thus, we took advantage of the temporal resolution of MEG to test specific predictions about the time course of activation: A region involved in referential search would be predicted to show an increase in activation as long as reference could not be resolved; on the other hand, a region involved in combining a referential expression with its referent (or processes enabled by this combination) would be predicted to show an increase when reference can be resolved upon reading the adjective.

Reference resolution on the adjective was associated with increased activation in left ventral prefrontal cortex, as well as bilateral ATL. Both effects had an early onset (between 200 and 300 ms after adjective onset), confirming that people interpret referential expressions incrementally. The direction of the effect suggests that this activation does not reflect a referential search process, but rather a process associated with being able to resolve reference. This could be building a model of ongoing discourse, but might also involve predictive processes (since the shape of the referent allowed predicting the subsequent noun) and anticipatory combinatory processes. In sum, our results provide MEG evidence for immediate reference resolution, implicating the VMPFC and ATL as regions relevant for this process.

References
[1] Eberhard, K. M. et al. (1995). *J Psycholinguist Res 24*, 409-36.
[2] Bemis, D. K. and Pylkkänen, L. (2011). *J Neurosci 31*, 2801-14.
[3] Nieuwland, M. S. et al. (2007). *NeuroImage 37*, 993-1004.
[4] Maris, E. & Oostenveld, R. (2007). *J Neurosci Methods 164*, 177-90.

Lexical processing and working memory in individuals with and without aphasia
Maria Ivanova (Center of Speech Pathology and Neurorehabilitation (CSPN)), Olga Dragoy (Research Institute of Psychiatry), Svetlana Kuptsova (CSPN), Anastasia Ulicheva (Hong Kong University), Anna Laurinavichyute (Higher School of Economics), & Lidia Petrova (CSPN)
mvimaria@gmail.com

Traditionally comprehension difficulties in non-fluent aphasia (NFA) have been regarded as syntactic in nature. However, individuals with NFA have been found to have considerable delay in lexical processing (Prather et al., 1997). In recent studies breakdowns in comprehension in NFA have also been explained as difficulties in lexical integration (Choy & Thompson, 2009, 2010). It remains unclear whether these emerging difficulties in lexical processing are primarily linguistic in nature or whether they are influenced by impairment of other cognitive processes, such as memory and attention (Friedmann & Gvion, 2003, 2012; Wright & Fergadiotis, 2012). The current study was designed to further tap into the time course of lexical processing in healthy controls and in NFA. We also aimed to investigate the potential role of working memory (WM) in efficiency of lexical integration.

To investigate lexical processing stories with temporally ambiguous sentences were employed in an eye-tracking-while-listening paradigm. All linguistic stimuli were presented in Russian. A story included two introductory sentences (1a), then a sentence with an ambiguous word (1b), and a disambiguation region (1c):

1a. *After the concert, the singer decided to have fun. First, he loosened his collar.*
1b. *Then he took a <u>shot</u> of*
1c. *icy vodka.*

In Russian two possible meanings of 'shot' include - 'a glass of vodka' and 'a pile of money'. The preceding context did not bias the interpretation of the ambiguous word prior to the disambiguation region. Stories were presented auditorily and were accompanied by visual panels with four drawings representing two meanings of an ambiguous word and two filler referents. Twenty experimental trials of this kind were presented. Participants' eye movements were recorded as they viewed the displays and listened to the stories. To measure WM the verbal eye-movement WM task (Ivanova & Hallowell, 2012) was used: participants had to visually match short and simple active sentences to one of the four pictures in the visual array and simultaneously remember a distinct set of colors for later recognition. Performance on the this task was monitored solely via eye movements. WM capacity was computed as the average proportion of fixation duration on target color sets in the recognition arrays. 56 native Russian speakers (36 healthy individuals, 20 individuals with NFA) participated in the experiment.

In the disambiguation region (1c) the mean proportion of fixation duration on the target (e.g. 'a glass of vodka') was significantly greater than on the competitor (e.g. 'a pile of money') and on the fillers for both controls (t (35) = 20.6, p < .001) and individuals with NFA (t (19) = 6.7, p < .001). However, for the NFA group the difference in proportion of fixation duration between the target and the competitor was significantly less than in healthy participants (t (54) = 5.0, p < .001), indicating less efficient lexical processing in individuals with NFA. Additionally, individuals with aphasia had reduced general verbal WM capacity compared to healthy controls (t (53) = 8.2, p < .001). A significant positive relationship between WM capacity and efficiency of ambiguity resolution was observed only for the aphasia group (r (17) = .45, p = .05).

The study demonstrated that both healthy controls and individuals with NFA resolve lexical ambiguity in the disambiguation region as soon as all the relevant linguistic information was provided. Though no significant delay in lexical integration of the correct meaning with the preceding context was observed for participants with NFA, overall, lexical ambiguity resolution was less effective in the aphasia group compared to healthy individuals. For the first time the link between efficiency of lexical processing and WM in NFA was demonstrated: our data suggest that general reduced verbal WM capacity places a limitation on resources available for lexical processing in NFA, making it less efficient than in healthy controls, possibly through insufficient inhibition of irrelevant meaning of the ambiguous word (Gadsby et al., 2008).

Neural correlates of sentence plausibility in garden-path processing
Dirk-Bart Den Ouden, Svetlana Malyutina & Victoria Sharpe (University of South Carolina)
denouden@sc.edu

Introduction. As part of a larger study into the neural correlates of cognitive operations that interact with or are indeed part of syntactic processing in sentence comprehension, we investigated *gradient* effects of plausibility on neural activation during garden-path (gp) processing. It was hypothesized that (re)analysis mechanisms induced by early-closure gp sentences are influenced by the plausibility of the gp interpretation. In these sentences, temporal ambiguity hinges on the initial interpretation of a postverbal noun as the direct object of the verb, rather than as the subject of a separate clause. Sentences in which the (syntactically erroneous) gp interpretation is plausible are expected to show the greatest neural activation in areas associated with thematic role association, while sentences in which this interpretation is made implausible possibly require less effort to restructure, but are expected to induce a response associated with general contextual processing.

Methods. An online survey was completed by 53 respondents, who rated the plausibility of the incorrect interpretation of gp sentences, in which the subject of the main clause was also assigned the role of the direct object of the subordinate clause. Respondents rated sentences such as "*The deer ran into the woods/paced in the zoo while the man hunted it*" on a scale of 1-7 (Question: "How plausible does the situation described in the sentence seem to you? Does it make sense, is it possible to imagine this happening?"). In an event-related fMRI experiment, 19 right-handed speakers of English (17 females, mean age 26, range 20-34) listened to sentences such as *While the man hunted the deer ran into the woods/paced in the zoo*, answering written comprehension questions by button press. GP sentences, non-gp late closure sentences, and non-gp fillers were presented with flat intonation as well as with normal intonation, i.e. with a natural pause after the embedded clause. Mean plausibility ratings were entered as parametric modulations of the BOLD responses to the sentence presentations at the first level of analysis. The resulting correlation maps were entered into a second-level t-test for the group analysis.

Results. Processing gp compared to non-gp (late closure) sentences elicited activation in inferior parietal and angular gyrus bilaterally, as well as right frontal areas. The opposite contrast elicited activation in left fusiform gyrus, bilateral medial frontal and anterior cingulate cortex, right sensorimotor cortex, and subcortical structures. Results show a positive correlation between BOLD response and global plausibility in bilateral superior temporal gyrus, with a left-hemisphere bias, extending into middle temporal and supramarginal gyrus. A negative correlation (or, rather, a correlation with implausibility), is observed in bilateral inferior parietal cortex, extending into angular gyrus and right-hemisphere supramarginal gyrus, as well as in bilateral frontal cortex, with a right lateralization.

Figure 1: Cortical areas that were positively (blue) and negatively (red) correlated with plausibility ratings.

Conclusion. This study aimed to tease apart neural activity associated with processing of temporally ambiguous sentences with varying levels of semantic plausibility. The left-lateralized posterior superior temporal activation associated with more plausible situations suggests greater effort in mapping thematic roles to syntactic structure for sentences that are more syntactically ambiguous, because the gp interpretation is plausible. There was considerable overlap between regions associated with processing of gp vs. late-closure (non-gp) sentences and regions where activation levels correlated with gp implausibility. As an implausible gp is assumed to require less restructuring than a plausible gp, the bilateral but right-lateralized inferior parietal and frontal activation pattern may reflect a general effort in making sense of the sentence as a whole, i.e., of the combination of the two clauses, rather than with syntactic restructuring *per se.*

Index

Index

A

B

C

D

E

F

Index

Index

Index