Hain't We Got a Right to use *Ain't* and Auxiliary Contraction?: Toward a History of Negation Variants in Appalachian English

Michael Montgomery
University of South Carolina

*Isn't, aren't, hasn't, haven't are strictly avoided. **Ain't** or **hain't**, depending on the need for emphasis, is used universally in the place of these "furrin" contractions. The bright mountain boy eager to improve his English but determined to avoid the use of **ain't**, is constrained to use **he's not** and **they've not** in lieu of **he isn't** and **they haven't** if he wishes to escape the damnatory label of being "stuck up."*
-Cratis D. Williams (1962: 16)

1. Setting the Scene

For quite a few years I have worked with the interviews that Joseph Sargent Hall recorded in and around the Smoky Mountains of Tennessee and North Carolina in 1939. Among other things, I have prepared a transcribed corpus of these short interviews (53,000+ words, 60 speakers all told), excerpts of which appear at a website called "Appalachian English", which was created in 2005 in connection with a dictionary Professor Hall and I co-edited (Montgomery and Hall 2004). [1] More recently, a dissertation student at Georgetown University contacted me seeking access to the corpus for the purpose of studying what has been called "McDavid's Law" (Troike 1986), a sound shift whereby [z] becomes [d] (both being co-articulated with glottal constriction or a glottal stop) before a nasal segment in the contractions *isn't*, *wasn't*, etc. The form *wasn't* was fairly plentiful (n=44) in the corpus, but *isn't* was not, and not because speakers usually framed their responses in the past tense. In unambiguous third-singular contexts for *be*, *isn't* occurred only three times, compared to nine instances of *ain't*, the dominant form. Two other variants were found, and these form part of the story to unfold later in this paper. Unfortunately, there were simply too few instances of *isn't* on

which to base any generalization and thus no point in sharing the Hall corpus with the disappointed student. I had become intrigued, however, about the prevalence of *ain't*, especially after finding that *aren't* never occurs in the corpus (vs. four instances of *ain't* and again two other variants) in third-plural and other contexts where *are* is usually found in mainstream American English.

Wolfram and Christian are doubtless correct in their assessment that "the usage of *ain't* is fairly extensive in [Appalachian English], perhaps more so than some other non-mainstream varieties of English" (1976: 116). As I looked deeper into the history and functions of *ain't* I found no documentation of it in the earlier English of Appalachia, very little for American and British varieties of English, and little in Britain generally. Establishing *ain't*'s course of development seems to have been a lingering problem in the philology of Modern English (Liberman 2006). The form had been examined on a quantitative basis by only a few modern sociolinguists (Feagin 1979, Wolfram and Christian 1976). Though its origin has been discussed and debated extensively, scholarship on its subsequent evolution seems hardly to have interested scholars. It is attested in England as early as 1695, and 18th century evidence is restricted very largely to the first-person singular (especially in tag questions), suggesting that one developmental path was the contraction of *am + not* to *amn't* to *an't*.[2] Except for scattered literary attestations, 19th century evidence is often indirect (as found in prescriptive grammars). In short, a good account has yet to be written detailing how and when the present-day form and functions of *ain't* have come to be and evolved. Other than for its early history as found in the literary record, little reliable historical evidence exists on its subsequent development on either side of the Atlantic.

The origin issue is threefold: 1) attestations are few; 2) because its etymology is obscure (i.e. English lacks a known element *ai-* to attach to the contraction of *not*;[3] and 3) *ain't* represents forms in the paradigms of two different verbs, *be* and *have*, what one linguist has termed a "diachronic coincidence" (Cheshire 1981:366). In perhaps the best-known effort to account for its origin and earliest history, Otto Jespersen posited multiple sources of *ain't*. Three of these were in the present tense of the *be*

paradigm (see 1), where the contraction of *not* with different verbal forms coalesced, according to Jespersen, to produce *ain't* through analogical formation in the 18th century (note: in early forms below and elsewhere, the apostrophe is inserted for convenience of discussion; it did not usually appear until the mid-19th century):

(1) *am not => amn't => an't => ain't*;
   is not => isn't => i'n't => an't => ain't;
   are not => aren't => an't/a'n't => ain't. (Jespersen 1961: 433)

Jespersen's proposals are problematic in being poorly attested and in sometimes relying on intermediate forms for which the orthographic evidence is very skimpy or that require phonological interpretations that are seemingly *ad hoc*. Further, he proposes a fourth source of *ain't*, from *han't* [i.e. a reduction of *haven't* (p. 431)], but he cites no example of this last form and does not relate it to the three others. If his account, with four different sources for *ain't*, strikes us as convoluted, it can be shown to be too simplistic on some grounds. However, his concern was to document and explore English prose and dramatic literature for the genesis and early use of the precursors of [ent], not with providing a unified account at any period. However important for the early groundwork it lays, his account offers few points that can be utilized for comparison (the only exception being his statement that "nowadays [ant] is frequently heard, especially in tag questions" [1961:432]). His perspective was that of the literary landscape in the 18th and early-19th century, with the interpretation of quasi-phonetic spellings made more uncertain by the incipient, scattered use of apostrophes. From his work one can conclude that [ent], spelled in various ways, was current in England by 1800 and represented forms in both the *be* paradigm and the *have* paradigm, but how current it was, socially and geographically, are unknown if not unknowable. One of the largest gaps in our knowledge about *ain't* left by Jespersen and other philologists is the trans-Atlantic one. Did *ain't* come to North America as a fully-fledged and functionally diverse form that it apparently was in England by 1800? What has been its history in American English? England certainly seems to have been the home of *ain't* in the British Isles. My research and many inquiries about

it in Scotland and Ireland, where it is not used, have consistently elicited judgments associating it with London or with urban southern England more generally. Even so, I will argue that one key to the wide geographical spread of *ain't* in the U.S. has been linguistic patterns brought by a stream of emigrants from Scotland and the north of Ireland (Ulster), numbering more than 150,000 people, to North America in the pre-Revolutionary 18th century. These emigrants did not bring *ain't*, nor did they bring *isn't* or *aren't*, but they did bring a type of negation that opened the door to other developments and to the spread of *ain't*, most particularly in Appalachia.

The scarcity of crucial data for *ain't* is of little mystery. *Ain't* has long had two strikes against its appearance in the written record. For one, it is a speech-based form. Despite meticulous surveys like Jespersen's, *ain't* shows up infrequently before the second quarter of the 19th century, when representations of vernacular speech, as in the work of Dickens, become decidedly more prevalent. For another, its social status in the U.S. has generally been not only colloquial but also stigmatized, so even in written representations of speech *ain't* is attributed only to certain character types or in only certain speech events. To supply sufficient valid examples to quantify in a meaningful fashion requires written evidence of a particular kind–from letters or other documents produced unselfconsciously by members of social classes having minimal formal education.

To contribute to filling in historical gaps and creating a fuller picture of the evolution of *ain't* in one region of the United States, the present investigation presents quantitative evidence from two corpora from Appalachia–one of written data and one of spoken–that are calculated to approximate vernacular speech insofar as possible. Using such corpora should bring us closer to answering questions about the development of *ain't* and forms that competed with it.

2. Background and Previous Study

According to Jespersen and other scholars, *ain't* arose as a negated form separately within two paradigms, *be* and *have*, in the English

of England by the end of the 18th century. Data from the *Corpus of Historical American English* (Davies 2012) reveal that by the early 19th century, *ain't* increased rapidly in American publications (n=5 for the 1820s and 326 for the 1830s), occurring as equivalents to negative forms of both *be* and *have*). Whether we consider them two functions or two lexemes, both have thrived, as *ain't* has been used in constructions like "she *ain't* here" and "she *ain't* left." In African American English, *ain't* also negates forms of *do*, particularly in the past tense ("she *ain't* go"; Green 2002: 173).[4] The general picture of *ain't* in American England is that of a widespread, socially conditioned form. *DARE* says nothing about the regional patterning of either function of *ain't* and states that *ain't* is "generally considered nonstandard" (s.v. *ain't* v[2]). *DARE* notes that the interrogative tag *ain't I* was documented for one-third of the cultured (i.e. educated, middle or upper-middle class) speakers interviewed by the Linguistic Atlas of the Middle and South Atlantic States in the 1930s and 40s (Atwood 1953: 31), but there was apparently no concentration of these speakers in any given sub-region.[5] Interestingly, *DARE* labels the form *hain't* as "chiefly NEast, South, South Midland, esp Appalachians." On such a note we turn to that last-mentioned region, from which *DARE*'s first unambiguous Appalachian citations of *ain't* and *hain't* are dated 1890 and 1886, respectively.[6] The research literature contains only one quantitative study from Appalachia useful for comparison, Wolfram and Christian (1976), based on sociolinguistic interviews conducted in two counties in southern West Virginia. Their volume provides the point of departure for later work on many linguistic features, including those for negation. Being interested largely in cross-generational comparison in order to assess the status of features in apparent time and thereby to posit change in progress, they profile the negation of *be* and *have* for 25 speakers (half or more of whom were apparently children aged 7 to 18) and find the usage of *ain't* to be stable across generations. Their combined results are shown in Table 1.

Table 1. *Ain't* in Southeastern West Virginia, not including first-person contexts (Wolfram and Christian 1976:116)

Be paradigm			*Have* paradigm		
ain't	*aren't/ain't*	*% ain't*	*ain't*	*hasn't/haven't*	*% ain't*
81	5	94.2%	49	23	68.1%

Wolfram and Christian conclude as follows:

> There are very few instances of the form[s] *isn't* and *aren't*; in fact, there are only four cases of *isn't* and one of *aren't*. For most [Appalachian English] speakers we can therefore conclude that *ain't* is almost categorically used instead of the standard English correspondences *isn't* and *aren't*. We note, by contrast, that the negativized contractions *haven't/hasn't* are far more variable in the speech of most of our [Appalachian English] sample, although *ain't* tends to be used more frequently than *haven't/hasn't*. (Wolfram and Christian 1976: 116)

Their figures of 94.2% *ain't* for negative *be* and 68.1% *ain't* for negative *have* are surprisingly high, but as will be shown, these are doubtless artificially high because of the way the investigators viewed negation. Wolfram and Christian make no suggestion for the infrequency of the "standard English correspondences," apparently attributing this finding to the fact that their speakers' English was rather "non-mainstream." Nor do they cite the frequency of *hain't* in their discussion of *ain't*, in relation to which they suggest elsewhere only that "the retention of initial *h* in these forms [i.e. *hit* and *hain't*] is most likely to be found in the more stressed items in a sentence, and occurs rarely, if at all, in unstressed items" (1976: 58). In the *Dictionary of Smoky Mountain English* (Montgomery and Hall 2004), I included a similar statement at the entries for *hit* and *hain't* because it reflected a long-held view that they occur primarily in sentence-initial or other stressed position ("*Hit*'s a-coming a rain," "*Hain't* it been a lousy day?") and that *hit* tends not to occur as the object of a verb, object of a preposition, etc.[7] Such a distribution of these two forms is what I expected before more recently scrutinizing linguistic corpora and finding more complex variation. I began to learn about

that complexity of *hit* from a colleague who observed that sometimes the form appears to mean "that" rather than "it" (Judith Bernstein, p.c. 2009). In other words, the pronoun can be deictic, even contrastive, in addition to being referential (these functions are communicated by relatively greater stress, usually realized by higher pitch). Thus, one finds sentences like "They's one more weed, but I don't know the name of *hit*." More refined observations about *hit* will go into the revised edition of the dictionary (Montgomery, Hall, and Heinmiller forthcoming). The foregoing comments about *hit* and *hain't* represent strictly a synchronic view, not necessarily anything historical.

3. The Present Study: Material and Scope

This study, the first to examine *ain't* in any variety of American English on a historical basis, aims to fill gaps in the history of this and the related form *hain't,* thereby helping to rewrite this history. In so doing, it examines the patterning of negative contraction more generally and proposes a novel syntagmatic account of how *ain't* spread historically. The material to be analyzed comes from two corpora. One is the *Corpus of American Civil War Letters* (*CACWL*), which Michael Ellis and I have been building since 2007 (Ellis and Montgomery 2011 *inter alia*). At this writing, *CACWL* comprises more than four million words, consisting of the transcription from manuscript of a handful of diaries and more than eight thousand letters written during the four years of the conflict by Civil War soldiers and their family members. The primary criterion for including documents in this corpus is the absence of punctuation, which is taken to indicate little formal schooling or reading experience, in other words, minimal literacy. In documents that lack punctuation one finds also spelling that is erratic (though usually phonetic to some degree) and grammar that is non-standard, all of which suggests that the individual was writing "by ear."[8] Some letters have unconventional spellings around thirty percent of the time, requiring one to read such documents aloud. Other than periods after initials or abbreviations, *CACWL* letters average perhaps 1.5 marks of punctuation. They are documents designed for hearing rather than silent reading. Ever growing, more

than a quarter of *CACWL* documents come from Northern states (especially Pennsylvania and New England), so ultimately the corpus will be used to study and map not only *ain't* but also many other features of American English.[9] The sub-corpus used here includes letters from parts of seven states in southern Appalachia, a body of material henceforth called the "*CACWL* Sub-Corpus." Perhaps half of the Civil War letters in it are from Western North Carolina, with much smaller numbers being from Eastern Kentucky, East Tennessee, Southwest Virginia, North Georgia, West Virginia, and Northeast Alabama. In many ways *CACWL* represents documentary evidence superior to any employed heretofore for reconstructing American English. A sketch of its advantages will be useful.

First, as outlined above, written documents produced by minimally literate writers, by virtue of their misspellings, erratic capitalization, and especially vernacular grammar, contain more valid, more direct evidence than other written material for American English. This evidence is much closer to speech than compilations of written material such as the 400-million-word *Corpus of Historical American English*, which are compilations of published texts, which are by definition, edited (Mark Davies, p.c. 2013). What's more, all *CACWL* writers have been tracked through the 1860 U.S. Census and military records, meaning that those in the sub-corpus were all ascertained to be natives of Appalachia. Extensive metadata is usually available on *CACWL* writers, especially their age and their home county.

Second, *CACWL* evidence pre-dates that of other large sources of speech-based material accessible. The writers who produced it were born largely from the late 1820s through the 1840s. In comparison to the scarce data available for *ain't*, referred to earlier, *CACWL* represents a huge assemblage of unedited manuscripts for linguists to analyze, almost certainly far more than all other written sources for vernacular American English combined from the 19th century, when many studies face the barrier that spoken evidence dates back only into the latter half of the 19th century (and then only by assuming apparent time).

Third, the *CACWL* Sub-Corpus of 498,313 words (approximately 70 writers) contains a robust number of tokens to

study, 125 of *ain't* (and its orthographic variants *aint* and *ant*) and 497 of *hain't* (and its variants *haint* and *hant*). It should be noted that in general, in Civil War times, untutored writers rarely employed apostrophes and that those in *CACWL* never did, further evidence of the lack of reading experience. The data at hand are sufficient to examine a wide range of comparative linguistic variation.

Fourth, the grammatical position for forms to be analyzed from the sub-corpus is fixed and readily identified, always immediately following the subject in declarative clauses. While phonological factors such as the nature of the preceding or following segment probably play a role in the rate of one type of contraction and the surface form of contraction, they are not relevant to the present study, which will focus on the choice of contraction strategy.

To be used for comparison and contrast are interviews from the Appalachian Oral History Project (*AOHP*) carried out through Alice Lloyd College in Eastern Kentucky in the early 1970s and on deposit at the college's library. This "*AOHP* Sub-Corpus" of interviews with 49 speakers comprising about 280,000 words was transcribed as part of an ongoing project involving and supervised by the author called the Archive of Traditional Appalachian Speech and Culture (ATASC).[10] For speech, these oral history recordings represent a relatively early body of material, with speakers born between 1880 and 1920. In the *AOHP* interviews, the grammatical position of the negated verbal form is not fixed. Several instances of *ain't* occur in tag questions such as *ain't they?* or *ain't it?* As Feagin points out (1979: 218), tag questions warrant a separate accounting. However, they need not be excluded from analysis because *ain't they* varies with *aren't they*, though not with *are they not*. The *AOHP* interviewed predominantly rural people who sometimes had grade-school education of as little as four or five years in school terms that were short by modern standards. These interviews contain a good number of tokens as well (*ain't* = 92, *hain't* = 21), not as many as in Civil War letters, but plenty of data for comparison. Given the much broader geographical catchment of the *CACWL* Sub-Corpus outlined above, the two corpora may not be altogether comparable in terms of region, but all the letters were from writers born in Southern Appalachia. The

AOHP speakers were born two to four generations later than the Civil War letter writers.

This study will leave issues of social and stylistic variation to other occasions, concentrating on the larger picture of negation, i.e. how *ain't* varies with other negation strategies for *be* and *have*. Most speakers of American English have *ain't* in their repertoire somewhere, even if they employ it only once every ten years in a burst of anger or in making a testy or jocular statement (one thinks of the common assertion, "you ain't seen nothing yet"). *Ain't* may be reserved for tag questions (*ain't it?*, *ain't I?*–as was the case for *LAMSAS* speakers), where it may serve as only a conventionalized phrase between sentences and be functionally similar to "you know," "don't you think?", "isn't that right?", etc. There are probably many sub-varieties of English distinguishable by where and when *ain't* is used. It is clear from the data to be examined here that *CACWL* writers and *AOHP* speakers had *ain't* as part of their core grammatical systems.

4. Issues of Negated Forms

To track and compare *ain't* and its competing negation strategies across time using corpora, we need to identify the range of possible variant forms. In this study we will consider five variants for both the *be* and the *have* paradigms.[11] Civil War letters may lack apostrophes, but they do often feature examples of the spellings *aint* and *haint*. They also have *ant* and *hant*. There are good reasons to interpret the latter two forms as representing pronunciations close or equivalent to [ent] and [hent]. Many varieties of American English have "split-*a*" vowel systems, in which /æ/ tenses and approaches /e/ before nasals (Hall 1942: 23–24, Labov 1994), with two phonetic forms participating differently in on-going phonological developments. In modern Appalachian speech, the negative form of *can't* is often pronounced like *cain't* [kent], a pronunciation that *DARE* labels as "especially in the South, South Midland." In the *CACWL* Sub-Corpus, *caint* occurs 29 times, while for the negative *be* the forms *aint* and *ant* each appears 56 times. Analogously, the historically low-central vowel of a word like *haunt* 'unseen spirit, ghost' variably had fronted to

[æ] in the 18th or 19th century, as attested in a Civil War letter from north Georgia by the spelling *hant*, as in (2), and the vowel has subsequently been raised.

> (2) Mother I have saw *hants* since I saw you tho that small particule of faith has led me through safe and sound. (W. P. Mangum, Jackson County, GA, 2 December 1862)

DARE characterizes the distribution of [hent] for *haunt* as "South, South Midland."[12] *AOHP* speakers from Eastern Kentucky use this pronunciation for *haunt* exclusively. Strictly speaking, the vowel of a given instance of *ant* or *hant* in Civil War letters is unknowable, but because [æ] was apparently moving toward or had reached [e], the spellings *ant* and *hant* in letters will here be considered evidence of the two variants *ain't* and *hain't*.

Examining the negation of *be* and *have* finds five potential variants in each paradigm. Three of them are *ain't*, *hain't*, and the full forms of the auxiliary + *not* (*is not, has not*, etc.). The fourth and fifth variants involve contraction strategies, of *not* (*isn't, hasn't*, etc.) or of the verbal form to the preceding word (*'s not, 've not*, etc.).[13] In conformity to other research (e.g. Yaeger-Dror et al. 2002), these strategies will be referred to as *not*-contraction and Aux-contraction. The following examples from the *CACWL* Sub-Corpus show nine of the ten possible manifestations of the negation strategies and provide a good taste of the language of Civil War letters written by minimally literate soldiers.

> (3) *Be* paradigm:
>
> a) *aint/ant*: [I] don't want to be intirley for gottan by you all if I *aint* a dooing you no good. (Isaac Copeland, Surry Co, NC, 26 February 1863)
>
> b) *haint/hant:* I *hant* no wais on easy a tall but what I will git home agin. (James Marshall, Forsyth Co, NC, 10 November 1862)
>
> c) *am*, *is*, *are* + *not*: I *am not* in the right state of mind to Write yours &c. (Larkin Kendrick, Cleveland Co, NC, 4 April 1862)
>
> d) *amn't/aren't/isn't*: if it *isnt* made in the Confederacy I dont know how we will get it. (W. L. Barrett, Pickens District, SC, 5 June 1863)

e) *'m not*/ *'re not*/ *'s not*: thair has Bin severil left hear with out fur lows and *its not* A going [to] stop. (9 July 1864) (John W. Reese, Buncombe Co, NC, February 1864)

(4) *Have* paradigm:

a) *aint/ant*: you must think that I *aint* got since enuff to know whether he is bad hurt or not. (Daniel Gilley, Henry Co, VA, December 1864)

b) *haint/hant*: I *haint* got narry letter from you in somtime. (Thomas Warrick, Coosa Co, AL, 23 June 1863)

c) *hav(e) not, has not*: I *hav not* mutch to Rite onely A Baut hard times. (John W. Reese, Buncombe Co, NC, 27 May 1863)

d) *hav(e)nt, hasnt*: I *havent* draud monny since June the 1st. (Chillon Carter, Kentucky, 14 September 1862)

e) [does not occur]

5. Evidence from Civil War Letters

Table 2 tabulates the five variants as they occur for the *be* paradigm in the *CACWL* Sub-Corpus (for illustrative purposes, *aint* and *ant* are separated as lines 1a and 1b). In this and subsequent tables data have been arranged into five columns by person and number of the subject and five (or sometimes six) rows for variant forms: *ain't, hain't*, full forms, *not*-contraction, and Aux-contraction. The discussion will be selective, focusing mainly on the rightmost column, with its subtotals by variant.

Table 2. Negated *be* Paradigm in *CACWL* Sub-Corpus

Variant	1st sing.	2nd sing./pl.	3rd sing.
1a *aint*	17 (14.9%)	2 (20%)	27 (12.4%)
1b *ant*	8 (7.0%)	1 (10%)	27 (12.4%)
2 *hant, haint*	6 (5.3%)	0 (--)	12 (5.5%)
3 *am, is, are + not*	83 (72.8%)	7 (70%)	148 (68.3%)
4 *amn't/aren't/isn't*	0 (--)	0 (--)	1 (0.5%)
5 *'m not*/ *'re not*/ *'s not*	0 (--)	0 (--)	2 (1.0%)
Total	114	10	217

Variant	1st pl.	3rd pl	Total
1a *aint*	3 (9.1%)	7 (17.5%)	56 (13.5%)
1b *ant*	1 (3.0%)	19 (47.5%)	56 (13.5%)
2 *hant, haint*	1 (3.0%)	2 (5.0%)	21 (5.2%)
3 *am, is, are + not*	28 (84.9%)	12 (30%)	279 (67.2%)
4 *amn't/aren't/isn't*	0 (--)	0 (--)	1 (0.2%)
5 *'m not/'re not/'s not*	0 (--)	0 (--)	2 (0.5%)
Total	33	40	415

In Civil War letters, *aint* and *ant* occur at the same overall frequency, with differences between the two in first-singular and third-plural contexts balancing each other out. *Aint/ant* occurs at a combined rate of 27.0%, *hain't* at 5.2%, and *not*-contraction at 0.2%. Aux-contraction occurs at a very low overall rate of 0.5% and is completely absent in the first-person singular. The only instance of *not*-contraction is seen as sentence (5):

> (5) we must have subsistence from some where & if it *isn't* made in the Confederacy I don't know where (W. L Barrett, Pickens District, SC, 5 June 1863)

The sentence in (6) is one of the two examples of Aux-contraction:

> (6) the Blue grass is A Bout halfe leg hy in the woods and feelds thats not tended (John R. Reese, Buncombe Co, NC, 23 April 1863)

Full forms are overwhelmingly dominant, with 67.2% overall and rates as high as 84.9% in first-person plural contexts. Only in the third-person plural are full forms surpassed by. Another variant, *aint/ant* (30% vs. 65%).[14] The rate of *aint/ant* is the second highest in all columns except in the third-plural. *Hain't* occurs a surprising 21 times, as in examples 7–9:

> (7) I think tha [=there] *hante* no danger but I will git home. (Hillary Shifflet, Madison Co, KY, 10 December 1861)

> (8) I *hant* very welll at present my old complaint is at work on me and has bin about a week. (J.C. Owens, Wilkes Co, NC, 26 April 1863)

> (9) if their is anny chance I am a gowing to come but if their *haint* I cant come. (William Tesh, Yadkin Co, NC, 2 Dec. 1862)

Haint/Hant, which seems oddly out of place in the *be* paradigm, is spread across person/number columns, everywhere except in the second person and at rates ranging between 3.0% and 5.5%. Examples do not occur in any fixed or formulaic phrase and are widely distributed among writers.

Like Table 1 showing Wolfram and Christian's findings from West Virginia, Table 2 demonstrates a great disproportion between *ain't* and *not*-contraction (in the *CACWL* Sub-Corpus the frequencies are 112 vs. 2). However, Table 2 and subsequent tables show that such a finding is only part of the picture of negated forms. We do not know whether Wolfram and Christian's speakers used full forms, Aux-contraction, or *hain't* (the last two being the other variants found in the Hall transcripts referred to at the beginning of this paper.) If Wolfram and Christian's speakers lacked full forms, this may have warranted limiting the counting to contracted forms, but in pursuing their primary objective of quantifying the occurrence of *ain't* across generations), they ignored or overlooked Aux-contracted forms.

More than two-thirds of negated *be* forms in the *CACWL* Sub-Corpus have *not* following a full form of the verb. This rate raises intriguing questions, such as whether the lack of *not*-contraction for *be* and *have* in the *CACWL* Sub-Corpus is consistent with a general paucity of *not*-contracted forms elsewhere or whether the use of full forms may reveal what Yaeger-Dror et al. (2002:80) call the Cognitive Principle, that writing promotes full forms, which carry "important semantic information" and are more salient. Perhaps it represents a writing tendency for minimally literate writers.

As we will see below, negative contractions in the letters are numerous, but they involve only certain forms. With so many misspellings and so many other indications of speech, not to mention their frequency of *ain't*, it is difficult to believe that soldiers patterned their writing after printed material (which probably had an artificially high rate of full forms). It is likewise difficult to imagine contractions not being very prevalent in the speech of the letter writers–but if so, why don't they show up? Whatever the reason, minimally literate writers from Appalachia in

the mid-19th century eschewed both *not*-contraction and Aux-contraction in writing. The only competitors to the dominance of negation by full forms of *be + not* were *ain't* and, to a far lesser degree, *hain't*. It would seem probable that writers did not perceive these two forms as contractions.

Table 3 presents a corresponding breakdown of negated forms in the *have* paradigm for the *CACWL* Sub-Corpus (with *hant* and *haint* separated for illustration). In part because *ain't* in this paradigm may, as argued by Jespersen, have a different origin than for the *be* paradigm, we might expect and indeed we find many intriguing contrasts to the *be* paradigm figures shown in Table 2. *Aint/ant* representing *have* or *has* is nearly non-existent, appearing only thirteen times in 1070 possible contexts (1.2%). Combined *haint/hant* occurs at the highest overall rate of 44.5% and is spread remarkably evenly across persons and numbers. Next comes full forms at 35.4%, a rate hardly half that for the *be* paradigm (67.2%).

Table 3. Negated *have* Paradigm in *CACWL* Sub-corpus

Variant	1st sing.	2nd sing./pl.	3rd sing.
1 *aint/ant*	5 (0.6%)	3 (4.2%)	3 (3.0%)
2a *hant*	268 (34.6%)	23 (31.9%)	34 (34.3%)
2b *haint*	74 (9.6%)	9 (12.5%)	9 (9.1%)
3 *hav*(e) *not, has not*	264 (34.1%)	20 (27.8%)	53 (53.5%)
4 *haven*(t)*, hasnt*	163 (21.1%)	17 (23.6%)	0 (--)
5 *'ve not/ 's not*	0 (--)	0 (--)	0 (--)
Total	774	72	99

Variant	1st pl.	3rd pl.	Total
1 *aint/ant*	0 (--)	2 (4.4%)	13 (1.2%)
2a *hant*	26 (32.5%)	20 (44.4%)	371 (34.7%)
2b *haint*	7 (8.8%)	6 (13.3%)	105 (9.8%)
3 *hav*(e) *not, has not*	33 (41.2%)	9 (20%)	379 (35.4%)
4 *haven*(t)*, hasnt*	14 (17.5%)	8 (8.9%)	202 (18.9%)
5 *'ve not/ 's not*	0 (--)	0 (--)	0 (--)
Total	80	45	1070

The complete absence of Aux-contraction may be in line with the *be* paradigm, but that of *not*-contraction is anything but that, with the latter contraction strategy being 0.2% for *be* vs. 18.9% for *have* (all of the 202 instances are *havnt* or *havent*; *hasnt* does not occur.) We see examples in (10) and (11):

    (10) I cant write a long letter paper is low and it is mity high and I havnt no money (Thomas Warrick, Coosa Co, AL, 18 April 1863)

    (11) we havent got no commissiond officer in the company Now (William Tesh, Yadkin Co, NC, 2 December 1862)

Occurring at barely half the rate for full forms, the relative frequency of *not*-contraction in the *have* paradigm implies the absence of a general constraint against contracted *not* among the letter writers, and it naturally leads one to ask which other verbal forms engage *not*-contraction. Table 4 presents for the *CACWL* Sub-Corpus the frequency and rate of a dozen forms that are contractible in present-day English.

Table 4. Contraction by Verb Form in *CACWL* Sub-Corpus

	Full form + *not*	Aux-Contraction	*not*-Contraction	*aint*/*haint*	Total
am	83	0	0	31	114
are	75	0	0	34	109
is	149	1	2	66	218
has	53	0	0	46	99
have	326	0	202	453	981
can	98	0	756[a]	--	854
could	102	0	1	0	103
will	111	0	242[b]	--	353
would	70	0	3	--	73
shall	4	0	13	--	17
do	113	0	1358	--	1471
do(e)s	10	0	0	--	10
did	15	0	6	--	21

[a] = includes 29 instances of *caint*.

[b] = a few cases may represent *want*.

In addition to *is* and *have*, we see that writers contracted *not* with several other verbs: *can*, *could*, *did*, *do*, *shall*, *will*, and *would*. The proportion of *cant* to *cannot* or *can not* and of *dont* to *do not* is more than 7:1. Examples are given in 12–20:

(12) you must look over mi bad wrighting and s[p]elling for i *cant* do nary won (Elizabeth Chapman, Campbell Co, TN, 29 May 1864).

(13) several of the boys went over to town yesterday to get salt & came back & said thay *couldnt* get mor (W. T. Martin, Pickens District, SC, 21 October 1861).

(14) the bulets and shel have sailed a Rowound me thick but not one *didnt* hit me at tall. (W. L. Barrett, Pickens District, SC, 24 July 1863).

(15) I *dont* think wee will stay here long ore at least I wood like to leave here (A. S. Harrill, Hancock Co, TN, 11 April 1864).

(16) I *havent* no news of interest to write at this time (Thorton Sexton, Ashe Co, NC, 12 November 1863).

(17) if it *isnt* made in the Confederacy I dont know how we will get it (W. L. Barrett, Pickens District, SC, 5 June 1863).

(18) I expect some hot times soon but I *Shant* fight much moore my Self (Hillary Shifflet, Madison Co, KY, 9 August 1862).

(19) I hope they *wont* fight eny more her for ther is no fun in it (Thomas Warrick, Coosa Co, AL, 13 January 1863).

(20) I cant wear it but I *wouldnt* take nothing in [the] world for it is a nice little ring (Molly Tesh, Yadkin Co, NC, 25 February 1864).

Altogether the *CACWL* Sub-Corpus features eight different verb forms taking *not*-contraction, in addition to *aint/ant* and *haint/hant*. The figures in Table 4 can be compared to ones from another corpus of "speech-like" texts, the Old Bailey Corpus, a compilation of trial proceedings from London's long-time central criminal court (Hüber 2009). The proceedings, recorded in shorthand at trials and then converted to long-hand and published, spanned the years 1674–1913. Hüber found that *can*, *will*, *shall*, and *do* were the only four verb forms that frequently took *not*-

contraction, whereas *is*, *did*, etc. did not). He attributes this finding to the fact that those four forms have modification of the vowel in the verb stems when they are negated. Essentially he argues that contracted *not* could not carry sufficient cognitive load by itself– that transcribers felt that either a full *not* or a change in vowel in recording trial discourse for negative statements in the testimony to be unambiguous. In other words, he posits that either an uncontracted *not* or a modified vowel was crucial to indicate the polarity of a statement when testimony could have major consequences for a defendant. The Old Bailey Corpus apparently lacks spellings like *caint*, so Hüber presumes that *can* and its contracted negated form had distinct vowels in colloquial English in England in the 1700s and 1800s. Hüber does not subject *have* to his test of vowel modification or report whether *have not* or *havent* was the more commonly negated form in the trial proceedings, and whether *havent* or another contracted form had a modified vowel is beyond the scope of this paper. Interestingly, he reports that the form *an't* (presumably representing *ain't*) did not show up before the turn of the 19th century.

The high rate of full forms (67.2%) seen in Table 2 and the lack of contraction for forms of *be* might by themselves support a claim that Civil War soldiers never used contractions, but the claim would be reasonable only if we considered the exponents of *be* in negated contexts. In fact, the writers use plenty of contractions elsewhere. They do not contract *not* with *are* or *has*, and they write *isn't* only once in nearly 500,000 words, but they contract with *have*, *do*, and *would*, and they write *cant* 756 times and *dont* more than 1300 times. Thus, the letters contain numerous contractions, but very rarely *is + not* to *isn't* and never *are + not* to *aren't*. The prevalence of full forms for these verbs is admittedly consistent Hüber's argument that either a vowel change or a full *not* was necessary to be unambiguous, but it hardly means that something else may not account for the lack of *not*-contracted forms.

It is interesting to glance at the *OED2* entry for *ain't*, which is still online at this writing (July 2014). Though it now has been replaced by a still-incomplete account in the online third version, *OED2* simply states (s.v. *ain't*) that *ain't* is a "[v]ariant of *hain't*," a statement that presumably means that the editors had in mind a

progression of *have not* to *haven't* to *han't* to *hain't* to *ain't*. This account is logical, attractive, and largely consistent with Table 3. However, the evidence from *CACWL* raises the obvious problem that, as of the middle of the 19th century, only 10.4% of the instances of *ain't* (13/125) were in the paradigm where they purportedly developed. The evidence in Tables 2 and 3 makes it difficult to argue that *ain't* had a single source, much less that the source developed from contracted *have*. The paucity of *ain't* in Table 3 suggests that, when representing negated *have*, the form was either a minor or an emerging variant. Of course, the writers of Civil War letters were born 165 or more years ago, and the *CACWL* Sub-Corpus gives us only a transitory, though generous, picture of English that was probably in transition, but transition from what and to what? Such questions provide as good of a reason as any to move to an examination of our second corpus.

6. Evidence from Oral History Interviews

As stated previously, the *AOHP* Sub-Corpus represents a subset of 49 interviews (approximately 280,000 words) from the larger Appalachian Oral History Collection of recordings at Alice Lloyd College. Two other people and I have transcribed these interviews as carefully as possible, putting them through a meticulous process of double auditing and correction, using a standard transcription protocol that calls for attending especially to the beginnings and endings of words in order to capture, among other things, forms like *hit* and *hain't*. Each of the transcribers spent all or most of their formative years in southern Appalachia. In the future one may need to consult the acoustic screen to ascertain whether aspiration is present to distinguish *ain't* from *hain't* more reliably, but when it comes to *not*-contraction and Aux-contraction, the main issue of transcription needing instrumental support would be to distinguish full forms from Aux-contraction.

Table 5 shows figures for the *be* paradigm for the *AOHP* Sub-Corpus. More than four out of five variants are either Aux-contraction (at 46.2%, n=72) or *ain't* (at 40.2%, n=63). *Hain't* comes in third place (at 9.0%, n=14), followed by *not*-contraction (at 4.5%, n=7). Examples are shown in 21–27:

(21) Hit*'s not* hurting me a bit in the world now.

(22) They*'re not* as sincere in their beliefs today as they used to be.

(23) They *ain't* going to be no such thing.

(24) They *ain't* nothing I can tell you about them.

(25) I *hain't* going in there and work today because my place has water in it.

(26) [There] *hain't* nobody knows. ("There is no one who knows.")

(27) That *isn't* going to hold it without it get set.

In the *AOHP* Sub-Corpus, the rates of *ain't* and *hain't* are clearly higher than in the *CACWL* Sub-Corpus (for *ain't* at 40.2% vs. 27% and for *hain't* at 9.0% vs. 5.2%), but the differences pale in contrast to the complete absence of full forms from the oral history interviews (vs. 67.2% for the *CACWL* Sub-Corpus). The dramatic contrast in full forms between the two data sources makes Table 5 look something like an inversion of Table 2.

Table 5. Negated *be* Paradigm in Eastern Kentucky (Based on Appalachian Oral History Collection, 1971–76)

Variant	1st sing.	2nd sing./pl.	3rd sing.
1 *ain't*	12 (27.3%)	3 (42.9%)	40 (48.8%)
2 *hain't*	2 (4.5%)	0 (--)	9 (10.9%)
3 *am, is, are + not*	0 (--)	0 (--)	0 (--)
4 *amn't/aren't/isn't*	0 (--)	0 (--)	6 (7.3%)
5 *'m/'re/'s + not*	30 (68.2%)	4 (57.1%)	27 (33.0%)
Total	44	7	82

Variant	1st pl.	3rd pl.	Total
1 *ain't*	2 (50%)	6 (31.6%)	63 (40.2%)
2 *hain't*	0 (--)	3 (15.8%)	14 (9.0%)
3 *am, is, are + not*	0 (--)	0 (--)	0 (--)
4 *amn't/aren't/isn't*	0 (--)	1 (5.3%)	7 (4.5%)
5 *'m/'re/'s + not*	2 (50%)	9 (47.4%)	72 (46.2%)
Total	4	19	156

Table 6 shows figures for the *have* paradigm in the *AOHP* Sub-Corpus interviews.

Table 6. Negated *have* Paradigm in Eastern Kentucky (Based on Appalachian Oral History Collection, 1971–76)

Variant	1st sing.	2nd sing./pl.	3rd sing.
1 *ain't*	15 (44.1%)	6 (42.9%)	4 (28.6%)
2 *hain't*	3 (8.8%)	1 (7.1%)	3 (21.3%)
3 *have not, has not*	0 (--)	1 (7.1%)	0 (--)
4 *haven't, hasn't*	10 (29.4%)	2 (14.4%)	4 (28.6%)
5 *'ve not, 's not*	6 (17.6%)	4 (28.6%)	3 (21.3%)
Total	34	14	14

Variant	1st pl.	3rd pl.	Total
1 *ain't*	3 (42.9%)	1 (20.0%)	29 (39.2%)
2 *hain't*	0 (32.5%)	0 (--)	7 (9.5%)
3 *have not, has not*	0 (--)	0 (--)	1 (1.4%)
4 *haven't, hasn't*	2 (28.6%)	2 (40.0%)	20 (27.0%)
5 *'ve not, 's not*	2 (28.6%)	2 (40.0%)	17 (23.0%)
Total	7	5	74

7. Comparing Findings from Two Appalachian Corpora

For the *have* paradigm in the *AOHP* Sub-Corpus, we observe *ain't* at an overall rate 39.2%, very close to that for the *be* paradigm rate (40.2%), but quite different from that for *have* in Civil War letters (1.2%) of earlier generations. Likewise, in the Eastern Kentucky oral history recordings, *hain't* and full forms approximate the rates in the *be* paradigm closely, at 9.0% vs. 9.5% for *hain't* and 1.4% vs. 0% for full forms. It is the two types of contraction that are so different between the two paradigms in the *AOHP* Sub-Corpus, with Aux-contraction for *be* having twice the rate for *have* (46.2% vs. 23.0%), but with *not*-contraction for *be* a tiny fraction of the rate for *have* (1.5% vs. 27%). The fact that *be* forms have initial vowels may be partially responsible for the different rates in the two paradigms. For both paradigms the two contraction variants constitute about half the tokens (77 for *be*, 37 for *have*), while *ain't/hain't* comprise the other half (79 for *be*, 37 for *have*). All in

all, the two paradigms for the oral history interviews look much more alike than do the two for Civil War writers.

Tables 7 and 8 provide summary comparisons for the two paradigms in the two corpora.

Table 7. Negation in *be* Paradigm Compared across Corpora

Variant	CACWL	AOHP
1 *ain't*	112 (31.4%)	63 (40.4%)
2 *hain't / hant*	21 (5.0%)	14 (9.0%)
3 *am not, is not, are not*	279 (67.2%)	0 (--)
4 *aren't, isn't*	1 (0.2%)	7 (4.5%)
5 *'ve not, 's not*	2 (0.5%)	72 (46.2%)
Total	415	156

Table 8. Negation in *have* Paradigm Compared across Corpora

Variant	CACWL	AOHP
1 *ain't*	13 (1.4%)	29 (39.2%)
2 *hain't / hant*	476 (44.5%)	7 (9.5%)
3 *have not, has not*	379 (35.4%)	1 (1.4%)
4 *haven't, hasn't*	202 (18.9%)	20 (27.0%)
5 *'ve not, 's not*	0 (--)	17 (23.0%)
Total	1070	74

Several points deserve highlighting. First is the striking prevalence of full forms in the *CACWL* Sub-Corpus, especially for the *be* paradigm, in comparison to the conversational speech of Eastern Kentuckians born two to four generations later (67.2% to 0%). The variants of negated *be* in the *AOHP* Sub-Corpus are almost exclusively Aux-contraction and (*h*)*ain't*. The *have* paradigm has a fair amount of *not*-contraction (18.9%), evidence consistent with the absence of a general prohibition against such contracted forms (Table 4), and this paradigm also presents a marked contrast between full forms (35.4%) and Aux-contracted ones (0%) in Civil War letters. The virtual absence of Aux-contraction for *be* in letters prompts one to ask how and whether it could have "surfaced" so quickly. It is possible, though unlikely, that Tables 7 and 8 reflect sub-regional differences to some extent.

The Eastern Kentucky speakers were a rural population whose immediate ancestors were principally also from the same area, and a large proportion of their previous ancestors came from Virginia, as did a high proportion of the fore-parents of *CACWL* writers.

It is also implausible that such marked differences between full forms and Aux-contraction are a product of dialect diffusion or contact whereby Eastern Kentuckians acquired Aux-contraction in the 19th century. It is true that timbering and then coal mining industries began arriving around the turn of the 20th century–when the Eastern Kentucky speakers were being born–but there is little evidence of contact with outsiders who used Aux-contraction. It is much more reasonable to posit that Civil War soldiers had Aux-contraction in their speech, but did not represent it in their writing. The *AOHP* Eastern Kentucky data, where full forms almost disappear (with only one token in the *have* paradigm), must reveal an approximation of how *CACWL* letter writers spoke. Comparing the two sets of data, we see that four variants other than full forms increase for the *be* paradigm, *ain't* by thirteen percent, *hain't* and *not*-contraction by four percent each, and Aux-contraction by 45.7%. For the *have* paradigm, three variants increase, most conspicuously *ain't* from 1.4% to 39.2% and Aux-contraction from 0% to 23%. The obvious presumption must be that full forms in Civil War letters disguise Aux-contraction and that phonologically the two were one and the same thing, which as we will see below one modern researcher (Feagin) realized when she sought to distinguish them on her recordings. Twentieth-century Eastern Kentuckians may have used full forms in deliberate, self-conscious speech, but such forms were evidently not part of their everyday language. Usually speakers contracted or used (*h*)*ain't*, raising the questions of where, if anywhere, they preferred *ain't* to Aux-contraction and where these two competing negation strategies came from.

A second point of interest concerns rates of *ain't*. While for the *be* paradigm it increased by nearly one-half (from 27.0% to 40.2%), in the *have* paradigm it jumped from 1.4% to 39.2%. Much, almost certainly most, of the latter shift can be attributed to the drop in *hain't* (44.5% to 9.5%). The figures suggest that *ain't* in the *have* paradigm was a relative late-comer in Appalachian

speech and provide further evidence that the developmental paths in the two paradigms were independent of one another. Such an assessment is not new, though lexicographers have had their difficulties in sorting it out. What we can glean from a quantitative study is how and especially when developments were taking place. Tables 2 and 7 show that in the *be* paradigm *ain't* was rather well established by the middle of the 19th century, though they suggest little about how it might have originated. *Aint/ant* for negated *be* was used in Civil War letters in all persons and numbers, as much as 65% of the time in the third-person plural. On the other hand, Tables 3 and 8, which break down negated *have* forms in Civil War letters, provide a veritable snapshot of the second path, from *have + not* undergoing *not*-contraction to *haven't*, then with [v] becoming a glide and vocalizing intervocalically to produce *hant* (a phonological process in English that has produced forms like *o'er* from *over* or *lord* from *hlaford* – the classic textbook example), then with the breaking of the vowel and the raising of the first vocalic element of the diphthong to *hain't*, and finally with the erosion of initial [h]. Evidence confirming what has been outlined here for *ain't* in Appalachia comes from the online *Corpus of Historical American English*, a compilation published texts from throughout the country. It contains no tokens of *ain't* in the 1810s and only five in the 1820s (all negated forms of *be*). In the 1830s the proportion of *ain't* representing negated *be* (35/336, or 10.4%) was exactly the same as for Civil War letter writers from Appalachia. Such a remarkable parallelism argues that the rise of *ain't* for negated *be* was an internal development for Appalachia.

Third, the data reveal an intriguing number of what we might call "cross-over" tokens of *hain't* in the *be* paradigm, 21 in the *CACWL* Sub-Corpus (e.g. sentences 3b, 7–9) and 14 in the *AOHP* Sub-Corpus (e.g. sentences 25–26). From the point of view of language processing and language change, it might seem most reasonable to consider such forms in Civil War letters as reversed spellings (Stephenson 1967) arising from the fact that as the sets of forms for the two paradigms had come to overlap, *ain't*, overlapped functionally in the minds of speakers, who then had difficulty keeping one set of forms apart from the other. As they used *ain't increasingly as a* form of negated *have*, speakers came

to use *ain't* and *hain't* in both paradigms and interchangeably. Such a phenomenon reveals both how language is learned and how language changes. As speakers manipulating the written code for the first time, Civil War soldiers were also learners of written English who, as writing researchers have shown in recent years, were often manipulating written English for the first time and in doing so manifest tendencies similar to a second language learner. That *hain't* occurs twice as often for negated *be* as for negated *have* in the *AOHP* Sub-Corpus (14 vs. 7) brings to mind another factor that might account for the proportions, which is prosody. Could the traditional, impressionistic claim that *hain't* occurs in more stressed, especially sentence-initial position, presumably irrespective of paradigm and with stress and aspiration being probably isomorphic, explain the fact that two-thirds of the instances of *hain't* occur in the *be* paradigm? In the *AOHP* Sub-Corpus only five of the 14 occurrences of *hain't* for negated *be* are found in sentence-initial position, and for the *have* paradigm only one of seven. The six sentence-initial instances of *hain't* are balanced by seven of *ain't* that occur there. All in all, prosody seems to be a poor predictor of *hain't*. The frequency (n=14) of *hain't* in the *be* paradigm might have been the errors of naive writers, but the form has become embedded in American English as a variant of negated, present-tense *be*. *DARE*'s entry (s.v. *hain't*[2]) features seventeen citations from 1887 to 1897, 9 of which come from their fieldwork in the 1960s, four of which are from the literary record of New England in the late-19th or early-20th century, and so on. It looks like *hain't* is here to stay.

Fourth, *not*-contraction appears to have been a peripheral strategy for negation in both corpora, at least in the *be* paradigm. As previous researchers have shown (e.g. Feagin 1979:218), in quantifying negative contraction it is a good idea to examine tag questions separately, because their frequency may skew sub-totals of *not*-contraction. Three of the six instances of *isn't* in the *AOHP* Sub-Corpus occur in the fixed phrase *isn't it* in the middle of a conversational turn. This phrase functions not as a yes-no question but as a discourse element to reinforce the speaker's fore-going point. The tag question is more or less equivalent to *you know*, as seen in (28):

(28) ... boys, both of them's got a little [baby], just walking
       around, ain't even married. Now that's right, *isn't it*?
       These little boys that's in school now, started them to raise
       a family and a-trying to provide for them. (July 13, 1972,
       interview with 57-year-old native of Floyd County)

Table 7 shows *not*-contraction to be the fourth highest (i.e. next to
least frequent) variant of negated *be* in the *AOHP* Sub-Corpus (at
n=7, 4.5%). Factoring out the three occurrences in tag questions
shows that the rate of *not*-contraction can be considered inflated
and that *not*-contraction is trivial in the *AOHP* Sub-Corpus.

8. Regional Perspectives

The lack of *not*-contraction in the *be* paradigm in Eastern
Kentucky is consistent with previous research. As cited above,
Wolfram and Christian document the low occurrence of *isn't* and
aren't in West Virginia, the lack of which made the rate of *ain't*
seem extraordinarily high. The same tendency for infrequency of
not-contraction characterizes varieties farther afield, as in
Anniston, Alabama, arguably on the southern edge of Appalachia
(Feagin 1979:218). Similarly, for Ocracoke Island, on the southern
end of the Outer Banks of North Carolina, several hundred miles
from Appalachia, Hazen (1996:102) found only a 4% rate of *isn't*
(4/99) and a 2% rate of *aren't* (1/50). His objective, like that of
Wolfram and Christian, was to contrast *not*-contraction with *ain't*,
but Feagin also sought to document the fuller scope of variation by
counting variants she identified as (*i*)*s no*(*t*), and (*a*)*re no*(*t*), which
combine what the present study refers to as "full forms" and "Aux-
contraction." Without stating as much, Feagin indicates that, based
on her listening, no line can effectively be drawn between the
contracted and unconstructed variants, a point that supports this
study's argument that full forms in Civil War letters disguise Aux-
contracted ones. Because of the initial vowel in forms of *be*,
probably few cases in Feagin's recorded interviews were
unambiguous full forms. Her results, summarized in Table 9,
closely approximate those from Eastern Kentucky shown in Table

7, except that she found no instances of *hain't* (Crawford Feagin, p.c. June 2014).

Table 9. Negation in *be* Paradigm for Anniston, Alabama Working Class (Feagin 1979: 219)

Variant	3rd singular	Plural	Total
ain't	111 (48.1%)	12 (30.8%)	123 (45.6%)
not-contraction	14 (6.1%)	2 (5.1%)	16 (5.9%)
Aux-contraction	106 (46.0%)	25 (64.1%)	131 (48.5%)
Total	231	39	270

The lack of *not*-contraction outside first-person singular, which she does not include, must be a wider-spread phenomenon than many scholars have realized.

Aux-contraction largely escaped the attention of researchers of regional American English until a ground-breaking, comprehensive study by Yaeger-Dror et al. (2002) seeking generalizations from the findings of a wide variety of corpora and producing a nationwide picture. Perhaps a tendency to view negated forms paradigmatically (seen in *not*-contraction) rather than syntagmatically accounts for the narrower perspective overlooking Aux-contraction. Yaeger-Dror et al. conclude that "the evidence shows that dialect area is a significant factor influencing *not*-contraction, with *not*-contraction preferred in the north (including New England, the northern cities area, the Ontario area, and the northwest), and Aux-contraction favored in the southeast and southwest ... For *is not* and *are not* the choice of one contraction strategy over another appears to be significantly correlated with the speaker's dialect area in conversation, adversarial, and literary situations" (2002:109). Interestingly, the researchers found insufficient *ain't* to quantify among their more than twenty collections drawn on, apparently none of which included extensive material from the rural United States.

When considered for the traditional speech of Eastern Kentucky, the preferred option for the contraction of *be* is clearly Aux-contraction. In the *AOHP* Sub-Corpus, speakers tend to say, "he*'s not* going," not "he *isn't* going." Likewise, Aux-contraction

with the preceding element is found for *are* ("They*'re not* coming"), *have* ("I*'ve not* seen my father"), *had* ("I wish he*'d not* made a glutton of himself"), *will* ("I*'ll not* speak his name"), and *would* ("She*'d not* leave that child").[15] Aux-contraction thus features a full complement of forms to *not*-contraction. It appears that in rural Appalachia, if Aux-contraction is not the first choice of negation strategies, it is usually second behind *ain't* for the *be* paradigm (as in the Joseph Hall recordings noted at the beginning of this study; cf. Montgomery forthcoming). Following Labov et al. (1968), Feagin proposes that speakers choose Aux-contraction in order to avoid *not*-contracted forms (especially *aren't* for *r*-less speakers) that resemble the stigmatized *ain't*. Such a strategy would seem highly doubtful in rural Appalachia. Who can blame Cratis Williams, a legendary scholar of the region's English, for calling *not*-contracted forms "'furrin' contractions" (see the prefatory quotation above)?

9. Historical and International Perspectives

According to the best evidence, *not*-contraction and Aux-contraction are first attested at about the same time, around 1600, in England.[16] It seems likely that they developed either sequentially or in different places and more reasonable to propose that Aux-contraction arose and spread first. In England this took place with *not* (a development from *nought* 'nothing'), and in Scotland with *nocht*, which evolved to *no*. In vernacular Scots today one finds sentences like "I'll *no* can help you," "He's *no* going," and "I'm *no* hungry." Later *nae*, a weak form of *no*, was enclitized, sometimes variably, to verb forms to produce *disnae* 'does not', *isnae* 'is not', etc., to produce a paradigm analogous to *not*-contraction. The dominance of Aux-contraction with the full form *no* over contraction with -n't or -nae is seen in Table 10, based on a study of middle-class, urban speakers in the 1970s (Miller 1980:6).

Table 10. Variation in Negative Forms in Scotland (Miller 1980:6)

	-n't/-nae	(*not-contraction*)
have	22 (59.5%)	15 (40.5%)
will	8 (42.1%)	11 (57.9%)
be	10 (9.0%)	101 (91.0%)

For these speakers, who like many of their peers in Scotland will have Anglicized their speech to one degree or another, we see rates favoring Aux-contraction that vary according to the form of the auxiliary, ranging from 91% for the *be* paradigm down to 42% for the *have* paradigm. Interestingly enough, for the *will*-paradigm the rate of Aux-contraction is 57% ("I*'ll not* tell you again") and is a pattern common in Ireland and Scotland today.

In gaining a view of the geography of negation strategies, it would not be wise to posit a dichotomy paralleling the Scottish-English border, because so much other linguistic variation does not. Instead, much of the northern half of England can be grouped with Scotland. Among others, Trudgill (1978) has proposed that the relative frequency of Aux-contraction increases the farther north one goes in Britain. Tagliamonte and Smith (2003) have sought to test such a claim by quantifying and comparing the contraction of *be*, *have*, and *will* in eight locations, two in Scotland, four in "northerly England," and two in Southern England. For the contraction of *be*, they find that Aux-contraction to be categorical in Scotland, but that in northern England contraction, the "choice of form can be explained by the influence of the preceding phonological environment." A far more extensive, if far less intensive, view of contraction comes from the findings of the *Survey of English Dialects* (Orton and Dieth 1962–71), as compiled by Upton et al. (1994: 495). This broad if not deep survey found *ain't* as far north as Yorkshire and Lancashire, where Aux-contraction was also documented. Historically and today *ain't* has not been a Scottish form, and when 20th-century linguists characterize it as a feature of "British English" (e.g. Cheshire 1981), this means "southern England." *Ain't* is no in Scottish dictionaries, and the closest to confirmation I have elicited from a

Scottish linguist for its use is that there's some evidence of *ain't* in Glasgow street language (Caroline Macafee, p.c. 1993).

The *Survey of English Dialects*, which documents *ain't* as quite widespread, is well known for targeting what have come to be called NORMs (non-mobile, older, rural, male) speakers, so, assuming apparent time, its evidence for *ain't* would seem to date well back into the 19th century. Other sources do not support an assessment. As mentioned earlier, the Old Bailey Corpus attests *ain't* (and its orthographical variants) only after the mid-point of the 19th century. The same can be said for the *English Dialect Dictionary* (Wright 1898–1905), although the latter source should be considered geographical rather than historical, because it did not always seek precursors for its late-19th-century material. Nor is there any evidence of *ain't* or its variants in what is the best-known sizable compilation of letters from English working-class emigrants (Erickson 1972). The letters transcribed by Erickson (100,000+ words) are similar to those in *CACWL* in having almost no punctuation, revealing numerous features of pronunciation, and attesting much non-standard grammar (Montgomery and Giner 1997), but there is no *ain't*. One must conclude that still very little is known about the currency of *ain't* and its competing negation patterns in 19th-century British.

10. Dialect Contact in America

Enough has now been established to posit a broad historical geographic picture, but only tentatively and generally. *Ain't* came to North America from England, as did *not*-contraction. It is probable that speakers from Scotland normally used Aux-contraction. In colonial days, the former group came largely in the 17th century and the latter in the 18th. The English were both earlier and far more numerous than Scots. However, tens of thousands of Scots had migrated to the north of Ireland in the 17th century, from where at least 150,000 people came to America in the pre-Revolutionary period 18th century. These descendants of Scots came to be known in America as the "Scotch-Irish," and they are known to have influenced American English heavily,

particularly putting a stamp on the language of the American interior, especially Appalachia (Montgomery 2004, 2006).

In short, I propose that Aux-contraction arose and spread in Britain first, followed by a competing pattern, *not*-contraction. The subsequent development of *ain't*, as well as *amn't to an't* to *ain't*, with i*sn't* perhaps going to *in't* and to *ain't*, as well as *amn't* to *an't* to *ain't*, took place only in England, especially southern England. At present we cannot estimate the prevalence of *ain't* during crucial North American migration and settlement times, leaving gaps for scholars to fill.

At least two, and perhaps three, vernacular patterns of contraction arrived in formative periods of American English. Dialect contact took place, at varying times and places, between speakers who brought Aux-contraction to America and speakers who brought *ain't* and *not*-contraction, leading to new dialect formation, as seen in the *AOHP* Sub-Corpus and in much of Appalachia today. It is clear from the history of the region and from existing scholarship that the two primary linguistic inputs were from English and Ulster Scots (Montgomery 1997, 2004). Speakers coming from Ireland or Scotland or having those linguistic heritages lived in close proximity with people from England or of English heritage. In many locales, not only in the American interior, one would expect a collision of negation patterns as communities of speakers formed.

Ain't, which can be seen as a suppletive form, has no doubt infiltrated the *be* paradigm at many times and places in North America where Aux-contraction was the dominant if not the only competing variant, risen to rates that cannot at present be estimated because of the lack of documentation for *ain't* and the contraction in general. Parallel developments have taken place when *not*-contraction has competed directly with *ain't* (as in parts of England and likely in New England) or with Aux-contraction (as in Scotland, as seen in Table 10 and likely in many parts of the American South). When variants having different sources compete, they can assume different values, as apparently has been the case for Aux-contraction and *not*-contraction in many of the corpora examined by Yaeger-Dror et al. (2002), with Aux-contraction tending to prevail in informative non-fiction because it permits the

full polarity marker *not*. Obviously *ain't* has the disadvantage of social stigma, this stigma being both the cause and the effect of legions of teachers, thousands of schoolrooms, and millions of textbooks and usage guides laboring to suppress it for the better part of two centuries. Ironically, such efforts may have imbued *ain't* with a salience and potency that is quite a different phenomenon from its relatively high rate of use in traditional rural communities like those cited earlier, where it may have escaped the efforts of educational forces. The reasons for its "escape" no doubt vary, but this fact does not make it any less important to consider why *ain't* has persisted as the most condemned word in English. At the same time, it is important to point out that, as an invariant form, *ain't* has had a distinct advantage in situations of new dialect formation, which that favor koinéization. As a hyper-form, not marked for person or number, *ain't* has the capacity to spread throughout a paradigm quickly and simultaneously replacing a variety of forms even in emerging varieties that have little or no previous use of *ain't*. Such a process has occurred in the formation of Anglophone pidgins and creoles (e.g. the rise of *ain* as the exclusive negating form in Gullah). In some varieties of English its spread may even have been aided by the form's use in the nearby paradigm of negated *have*.

11. Concluding Discussion

In part for the sake of argument, in part to unpack the earlier history of varieties of English in Appalachia, I have drawn an overly simple, over-arching distinction between two linguistic streams that came to North America in the 17th and 18th centuries. Scotland and Ireland may be areas peripheral to England today in population and power and wealth and seemingly language. But in terms of the number of immigrants that came to colonial America and particularly in terms of where many of their descendants went, we can discover a different story. Even so, I do not propose that the two linguistic streams can be taken simply to represent streams of population, because migrants, especially groups of migrants, usually have access to multiple varieties. It may not be possible to posit a similar scenario or set of linguistic dynamics for

grammatical sub-systems other than the negation of *be* and *have*, which involve multiple variants and the availability of a hyper-form like *ain't* that can resolve conflicts between Aux-contraction and *not*-contraction.

One advantage of the account presented here is its logical sequence for the development of contraction, the emergence of *ain't*, and its success in competing with other negating forms in vernacular English, not only in Appalachia, where it reached near-party with Aux-contraction. It proposes that Aux-contraction occurred first. This perhaps took place because forms like *am*, *is*, and *are* begin with a vowel, and it can take little time for such forms to encliticize to their subjects.

Another advantage of my proposals is that they offer an alternative to an all-encompassing, unitary history of *ain't* consistent with what Jespersen argued. I argue that *ain't* did arise in England at least once, but because of the competition between Aux-contraction and *not*-contraction, it may well have spread through dialect diffusion ultimately from England or dialect contact in North America. The development of *ain't* from *hain't* is something that could happen synchronically at any time, because the phonological process involved is a common one. In Appalachia the development may have taken place as late as the mid-19th century. So there is no reason to believe that *hain't* hasn't been created at several times or places from what was ultimately a form of *have not*, contracted to *haven't*. The present proposal need not rely on a single point of origin of *hain't*.

Before being approached by the Georgetown student, I would have vowed that *isn't* was a very common form. But in 500,000 words of Civil War letters and nearly 300,000 words of traditional speech from Appalachia it and other forms of *not*-contraction can hardly be found. Why have researchers of Appalachian English been so blind to Aux-contraction? Perhaps one is our tendency to think of verbs in terms of columns of paradigmatic forms. Only in looking from left to right does one see how much Aux-contraction there is. In looking in both directions we realize all five different variants and that for negation almost certainly dialect contact processes like leveling took place in North America in the 18th century and in Appalachia in the 19th and 20th. It may still be

taking place today. The English of Appalachia has long been reputed and is still widely viewed as a collection of relatively static varieties of American English. However, what evidence from corpora and a reconstruction of the history of negation have begun to suggest is that it may be as dynamic as any other.

NOTES

[1] See http://artsandsciences.sc.edu/engl/dictionary.

[2] Here and elsewhere *ain't* is used as a cover form for the orthographic variants used in the 18th and especially 19th century. With regard to the currency of *ain't* in the 18th century, I queried Terttu Nevalainen, Director of the University of Helsinki's *Corpus of Early English Correspondence* project. She checked a part of the *CEEC-Extension* section of the corpus, sending the writer three citations, all in first-person singular, and reported that she had "reason to expect that 'ain't' is very rare in the corpus ... None of us seem to have come across any [citations] in the earlier centuries" (Terttu Nevalainen, 11 November 2009).

[3] Perhaps the only exception to this was the intrepid Noah Webster, who, in his *American Dictionary of the English Language* (1828), stated that *ai* came from *er*, the Danish form of *be*.

[4] There is some evidence that *ain't* can function as a past tense in the English of Appalachia. Consider the following examples from the *AOHP* Sub-Corpus, analyzed in this paper:

> [If] they would uh, would [have] needed any kind of help a-tall, a-tall, the neighbors around would come in, you know, and they was just, hit was free, and they *ain't*, no-, nobody back then charged nothing. (M. Martin, 1971).
>
> It *ain't* many years ago, either. (E. Salmons, 1974).

Also, in her study in Anniston, Alabama, Feagin recorded two unambiguous examples.

[5] In surveying records for the Linguistic Atlas of New England and the Linguistic Atlas of the Middle and South Atlantic States, Atwood provides summaries of variants for *am not* and *have not*, but not *is not*.

[6] *DARE* includes an 1835 citation of *an't* from David Crockett's *Account of Col. Crockett's Tour to the North and down East, in the Year of Our Lord One Thousand Eight Hundred and Thirty-Four*. Philadelphia: Carey and Hart. However, this volume is regarded by Crockett scholars as likely to have been ghost-written.

[7] For an early exposition of this view, see Allison (1929).

[8] This does not mean that in their spelling writers do not sometimes employ what at first glance might appear to be hypercorrections, but that here these are taken to be reversed spellings, revealing variation that may also represent change in progress (as *mild* for *mile*, showing variation in final consonant

clusters). However, forms that imply the influence of writing or formal speech (as *whom* for *who*) on grammar are conspicuously absent. Nor does it mean that writers do not often employ formal phrasing in the opening or closing of their letters.

[9] A preliminary view of *CACWL* writers from New England (M. Ellis, June 2014) reveals much more not-contraction than in Appalachia (e.g. n=22 for *isn't*, n=12 for *hasn't*) and that *hain't* is used for *have/has not* 75% of the time (vs. 10% in Appalachia).

[10] ATASC is a compilation of transcribed oral history interviews from ten different areas of Southern Appalachia.

[11] Variants with negation by *no* (e.g. "There is no telling when it will stop") are not included in the present study because they involve only contexts in which a following nominal is indefinite.

[12] According to the Linguistic Atlas of the Middle and South Atlantic States, [æ] was widespread among older speakers interviewed in the 1930s. *LAMSAS* also found that "*haunted*, with the vowel /e/ of *eight*, occurs in parts of the Carolinas, Georgia, and West Virginia" (Kurath and McDavid 1961: 161).

[13] Both auxiliary and main verb forms of *be* and *have* are included.

[14] Future research may profit from examining variants according to subject type.

[15] These examples are taken from Montgomery et al. forthcoming. The only study that compares the rate of Aux-contraction by verb form in Appalachia is Nicholas (1977). Using a structured written elicitation, he ranked verbs as follows with respect to Aux-contraction: *are* (30.7%), *is* (28.9%), *will* (23.4%), *have* (9.9%), *has* (7.1%), and *would* (4.8%).

[16] Jespersen states that "the contracted forms [of *not*] seem to have come into speech ... about the year 1600," a conclusion reached from the meter of poetry having forms such as *wonnot* 'will not' (1966: 117). For a different view, see Rissanen (1999). Barber finds that in print *n't* "is not recorded until the middle of the 17th century" and is "rare until late in the century" (1976: 254). In perhaps the most detailed study, Brainerd finds the earliest occurrence of *n't* in 1621 and cites an example of *han't* (from Richard Brome's play *The Sparagas Garden* [1635]) that resembles a precursor to *hain't*.

REFERENCES

ALLISON, VERNON C. 1929. On the Ozark pronunciation of "it." American Speech 4: 205-06.

ATWOOD, E. BAGBY. l953. A Survey of Verb Forms in the Eastern United States. Ann Arbor: University of Michigan Press.

BARBER, CHARLES. 1976. Early Modern English. London: Deutsch.

BRAINERD, B. 1989. The contraction of *not*: A historical note.

Journal of English Linguistics 22: 176–96.

CHESHIRE, JENNY, Jenny. 1981. Variation in the use of *ain't* in an urban British dialect. Language in Society 10: 365–81.

ELLIS, MICHAEL E. and MICHAEL MONTGOMERY. 2011. About *all*: Studies in nineteenth-century American English I. American Speech 86: 340–54.

ERICKSON, CHARLOTTE. 1972. Invisible Immigrants: The Adaptation of English and Scottish Immigrants in Nineteenth-Century America. Ithaca, NY: Cornell University Press.

FEAGIN, CRAWFORD. 1979. Variation and Change in Alabama English: A Sociolinguistic Study of the White Community. Washington, DC: Georgetown University Press.

GREEN, LISA J. 2002. African American English: A Linguistic Introduction. Cambridge: Cambridge University Press.

HALL, JOSEPH S. 1942. The Phonetics of Great Smoky Mountain English. New York: King's Crown Press.

HAZEN, KIRK. 1996. Linguistic preference and prescriptive dictum: On the phonological and morphological justification of *ain't*. Sociolinguistic Variation Data, Theory, and Analysis: Selected Papers from NWAV 23 at Stanford, 101–12. Stanford, CA: Stanford University Department of Linguistics.

HÜBER, MAGNUS. 2010. Trial proceedings as a source of spoken English: A critical examination based on negative contraction in the Proceedings of the Old Bailey, 1674–1913. Anglistentag 2009, ed. by Jörg Helbitö and Réne Schalleberger, 65–78. Trier: Wissenschaftlicher Verlag Trier.

JESPERSEN, OTTO. 1940. General history of negated sentences. Modern English Grammar on Historical Principles: Part V (Fourth Volume), 426–39. London: Allen and Unwin.

LABOV, WILLIAM. 1994. Principles of Linguistic Change: Volume 1, Internal Factors. Cambridge, MA: Blackwell.

LABOV, WILLIAM, PAUL COHEN, CLARENCE ROBINS, and JOHN LEWIS. 1968. A Study of the Non-Standard English of Negro and Puerto Rican Speakers in New York City. Volume I: Phonological and Grammatical Analysis. Final Report. Washington, DC: Office of Education Cooperative Research Project No. 3288.

LIBERMAN, ANATOLY. 2006. The much vilified *ain't*. OUP Blog. http://blog.oup.com/2006/08/the_much_vilifi/.

MILLER, JIM. 1980. Negatives in Scottish English. Social Science Research Council Working Paper.

MONTGOMERY, MICHAEL. 1997. The Scotch-Irish influence on Appalachian English: How broad? How deep? Ulster and North America: Transatlantic Perspectives on the Scotch-Irish, ed. by Curtis Wood and H. Tyler Blethen, 189–212. Tuscaloosa: University of Alabama Press.

_____. 2004. Solving Kurath's puzzle: Establishing the antecedents of the American Midland dialect region. The Legacy of Colonial English: The Study of Transported Dialects, ed. by Raymond Hickey, 410–25. Cambridge: Cambridge University Press.

_____. 2006. From Ulster to America. Belfast: Ulster Historical Foundation.

_____. (forthcoming). Archive of Traditional Appalachian Speech and Culture. Typescript.

MONTGOMERY, MICHAEL and MARÍA GARCÍA BERMEJO GINER. 1997. Regional British English from the nineteenth century: Evidence from emigrant letters. Current Methods in Dialectology, ed. by Alan S. Thomas, 167–83. University of Wales.

MONTGOMERY, MICHAEL, JOSEPH S. HALL, and JENNIFER N. HEINMILLER, EDS. Dictionary of Southern Appalachian English. Knoxville: University of Tennessee Press.

NICHOLAS, J. KARL. 1977. I'll not say I won't do it. Paper presented at the Fall meeting of the Southeastern Conference on Linguistics.

ORTON, HAROLD and EUGEN DIETH. 1962–1971. Survey of English Dialects. Leeds: University of Leeds.

RISSANEN, MATTI. 1999. *Ain't it* or Is it not?: On the order of postverbal subject and negative particle in the history of English. Negation in the History of English, ed. by Ingrid Tieken-Boon van Ostade et al., 189–200. Berlin: Mouton de Gruyter.

STEPHENSON, EDWARD. 1967. On the interpretation of occasional spellings. Publication of the American Dialect Society

48:33-50.

TAGLIAMONTE, SALI and JENNIFER SMITH. 2003. Either it isn't or it's not: *neg/aux* contraction in British dialects. English World-Wide 23 (2): 251–281.

TROIKE, RUDOLPH. 1986. McDavid's law. Journal of English Linguistics 19: 177–205.

TRUDGILL, PETER. 1978. Sociolinguistic Patterns in British English. Cambridge: Cambridge University Press.

UPTON, CLIVE, DAVID PARRY, and J. D. A. WIDDOWSON. 1994. Survey of English Dialects: The Dictionary and Grammar. London: Routledge.

WILLIAMS, CRATIS D. 1962. Verbs in mountain speech. Mountain Life and Work 38.

WOLFRAM, WALT and DONNA CHRISTIAN. 1976. Appalachian Speech. Arlington, VA: Center for Applied Linguistics.

WRIGHT, JOSEPH, ed. 1898–1905. The English Dialect Dictionary. 6 volumes. London: Frowde.

YAEGER-DROR, LAUREN HALL-LEW, and SHARON DECKERT. 2002. *It's not* or *isn't it*? Using large corpora to determine the influence on contraction strategies. Language Variation and Change 14: 79–118.

Michael Montgomery
Professor Emeritus
University of South Carolina
Department of English Language and Literature
[ullans2000@gmail.com]