Locating Linguistic Variation: A Case Study of English Auxiliary Contraction

Laurel MacKenzie

University of Pennsylvania

1.   Introduction

It was noted in the early days of sociolinguistics that the grammatical locus of some variable phenomena — specifically, those that implicate linguistic levels above the phonology — may be difficult to pinpoint in a derivational architecture. Thus, Labov (1969, 721), in his seminal study of variable copula deletion in African American English, asks:

> By what kind of rule are these finite forms of *be* deleted? Is it a transformational rule which deletes the copula, or a separate set of rules which delete *is* and *are*? Or is it a phonological rule which operates at a lower level in the grammar?

Labov goes on to argue that the observed patterning of deletion with regards to other variable phenomena (specifically, copula contraction) argues for one analysis in particular: that by which forms of *be* are deleted by phonological processes. The more broadly-applicable argument is that the patterns and correlations observed in natural speech data provide evidence for the particular locus of this variable.

This paper seeks to revisit the bigger question that Labov raises: how can the locus of variation be identified when a variable process implicates many linguistic levels? I approach this question by means of a case study of English auxiliary contraction, a complex but variable process that, like copula deletion, could in principle have a number of plausible grammatical loci. I argue that previous work on the grammatical mechanism behind this variable has missed crucial evidence for its locus that can be obtained only through a quantitative study of natural speech.

This paper, then, has two main goals. The first is to supplement our knowledge of English auxiliary contraction via a large-scale corpus study that documents the surface patterns displayed by a number of contractable auxiliaries. The second goal is to revisit Labov's question of the locus of linguistic variation by developing a model of auxiliary contraction that accounts for the patterns evident in natural speech. Specifically, I put forth

a model which localizes variation at two levels of a derivation: a variable morphosyntactic process governs an alternation between full and contracted allomorphs; then, subsequent low-level phonological processes act on those allomorphs.

2.   Background
2.1   Overview of the Phenomenon

The phenomenon under study in this paper is the variable contraction of English auxiliaries (e.g. *He <u>had</u>* ~ *he'<u>d</u> been there all day*). The auxiliaries that can be contracted in Standard English — in all cases, leaving their final consonant — are *had, has, have, will, would, is, are, am, does,* and *did* (these last two only in *wh*-questions). The contraction of the first six of these auxiliaries will be examined in this paper.

2.2   Previous Work

The large body of theoretical literature on contraction has proposed a number of different grammatical loci. The earliest works (e.g. Zwicky 1970) analyzed contraction as the output of a cascade of phonological processes. One such process deletes an auxiliary's initial *h* or *w*; Zwicky attributes this to the same process that deletes the initial *h* of pronouns such as *he* or *himself* in fast speech. Another deletes the auxiliary's vowel, leaving a single consonant which forms a syllable with its host.

Kaisse (1985), on the other hand, argues against a purely phonological analysis, on the grounds that deletion processes in English are restricted to fast speech, while contraction is acceptable at any speaking rate. She provides an alternative analysis under which full and contracted forms are distinct allomorphs, rather than the latter being derived from the former. Under Kaisse's analysis, contraction is a variable morphosyntactic process adjoining an auxiliary to its host. A host-adjoined auxiliary then conditions the insertion of the contracted allomorph. The process is roughly sketched in (1) with the future auxiliary *will*; "X" represents the final item in a noun phrase. While the exact mechanics of host-auxiliary adjunction have received a number of different treatments in the literature, most recent analyses (e.g. Inkelas and Zec 1993, Wilder 1997) have at least upheld this account of full and contracted forms as allomorphs. I will adopt this analysis of contraction as morphosyntactic adjunction conditioning insertion of a contracted allomorph, assuming also that there is some further process of phonological cliticization (henceforth, "syllabification") that incorporates the contracted allomorph into a syllable with its host.

(1)   Adjunction feeding insertion of contracted allomorph
        $X^\frown T_{[fut]} \rightarrow [X[T_{[fut]}]]$ X and T form a complex head when adjacent
        $T_{[fut]} \leftrightarrow$ *'ll* / $[X[\_\_]]$ T is spelled out as contracted when in a complex head
        $T_{[fut]} \leftrightarrow$ *will*   T is spelled out as a full form otherwise

One shortcoming of the theoretical work on contraction is that it has tended to overlook the fact that the process is variable, and that the factors conditioning that variability may have implications for how contraction must be situated in the grammar. For instance,

if contraction is found to be conditioned by surrounding segments, the most direct interpretation of this would be that it occurs at a stage of the grammar at which phonological segments are visible; if it shows an effect of auxiliary identity, it must occur at a stage of the grammar at which the feature content of an auxiliary is visible. Quantitative data, which was lacking from those early studies, is necessary for determining what conditions contraction. In fact, quantitative studies of contraction, beginning with that of Labov (1969), have been carried out by a number of sociolinguists. However, they have almost exclusively focused on the contraction of *is* and *are*; in the rare event that the contraction of other auxiliaries has been examined — namely, McElhinny (1993) — the full range of auxiliary realizations has not been taken into account.

3. Methodology of the Present Corpus Study
3.1 Corpus

4898 tokens of the auxiliaries *had, has, have, is, will,* and *would* were pulled at random from the Switchboard corpus (Godfrey et al. 1992), a corpus of telephone conversations between strangers on a given topic ($\sim$240 hours/3 million words long). Tokens in the data set were uttered by 464 unique speakers. Tokens were located using transcripts, but all were hand-coded based on audio.

3.2 Defining the Envelope of Variation

Preliminary analysis of the data revealed that auxiliaries surface in three distinct forms. A coding scheme was thus adopted to capture these three forms, as follows:

1. Full: He [hæd]/[həd] been there all day.
   (initial consonant, audible vowel of any quality)
2. Intermediate: He [əd] been there all day.
   (no initial consonant but audible vowel)[1]
3. Contracted: He[d] been there all day.
   (no initial consonant, no vowel; final consonant forms a single syllable with its host)

Much work has focused on the environments in which contracted forms are blocked from occurring. This blocking can be due to the presence of a gap to the right of the auxiliary (the subject of much work, beginning with King 1970), or to the quality of the material to the auxiliary's left (see, e.g., Kaisse 1983, Sells 1983). Because there are environments in which auxiliary realization is said to be categorical, then, care was taken to omit these environments from the study, as they are external to the envelope of variation. As a result, tokens in which an auxiliary preceded a movement (2) or a deletion (3) site, or appeared in a comparative sub-deletion construction (4), were not coded. Forms of *is* were also omitted in pseudo-cleft constructions (5). Instances of preposed constituents, which are said by Kaisse (1983) and Sells (1983) to block contraction (6), did not occur in the corpus data.

---

[1]In the case of *is*, no intermediate form distinct from the full form can be distinguished; all uncontracted forms of *is* were coded as full.

(2)   I wonder where Gerard {is / *'s} __ today.   (King 1970)

(3)   I'm as tall as Bill {is / *'s} __.   (Sells 1983)

(4)   Marie's a better scientist than Pierre {is / *'s} an __ engineer.   (Anderson 2008)

(5)   What I wonder {is / *'s} whether we'll win.   (Kaisse 1983)

(6)   Speaking for Jakobson {is / *'s} Professor Chomsky.   (Sells 1983)

What these omitted environments have in common is that they all categorically allow one single allomorph; as a result, they are inappropriate for a study of variation. On the other hand, there are some environments in which contracted forms are illicit, but variation — between full and intermediate forms — still occurs. I am referring specifically to the auxiliaries *had, have, will,* and *would* when they surface after non-pronoun subjects (7–10).

(7)   Marie + *had* → [məˈri hæd] ∼ [məˈri əd], *[məˈrid]

(8)   three + *have* → [θri hæv] ∼ [θri əv], *[θriv]

(9)   Sue + *will* → [su wɪl] ∼ [su əl], *[sul]

(10)   Marie + *would* → [məˈri wʊd] ∼ [məˈri əd], *[mərid]

Because environments like those in (7–10) permit variation in auxiliary realization — even though that variation does not include the surface contracted form — such environments *were* retained as being within the envelope of variation.

4.   Findings

Previous work on contraction of the copula (e.g. Labov 1969) has found that contraction shows a strong effect of subject type: specifically, contracted forms of the copula are more frequent after a pronoun than after a full noun phrase (i.e. non-pronoun) subject. For that reason, auxiliaries in the present study were likewise coded for which type of subject they followed. For the purposes of this analysis, personal pronouns (henceforth "pronouns") were opposed to non-pronoun (henceforth "noun phrase") subjects. Quantifiers, demonstratives, the expletive pronouns *it* and *there*, and single *wh*-words were not included in either group, but multi-word *wh*-phrases, e.g. *What city {is / 's} she from?*, were retained in the "non-pronoun" category.

Auxiliary realization in post-pronoun and post-noun phrase environments is plotted in Figures 1 and 2. In Figure 1, personal pronouns have been restricted to those ending in vowels (which eliminates only *it*) where the auxiliaries *had, have, would,* and *will* are examined. This was done in order to exclude cases in which an auxiliary was barred from close attachment to its host for phonotactic reasons: e.g. *it'll* [ɪɾəl], *it'd* [ɪɾəd], in which the contracted form is unsyllabifiable with the preceding /t/. Tokens of *is* and *has* after *it* are nevertheless included in the plot, because contraction in those cases is phonotactically licit. In essence, then, the envelope of variation in Figure 1 has been restricted to only those cases in which contraction is phonotactically licit.
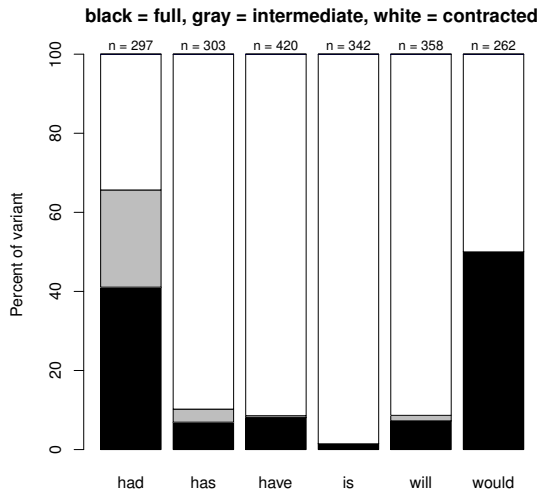
black = full, gray = intermediate, white = contracted

Figure 1: After pronoun subjects.

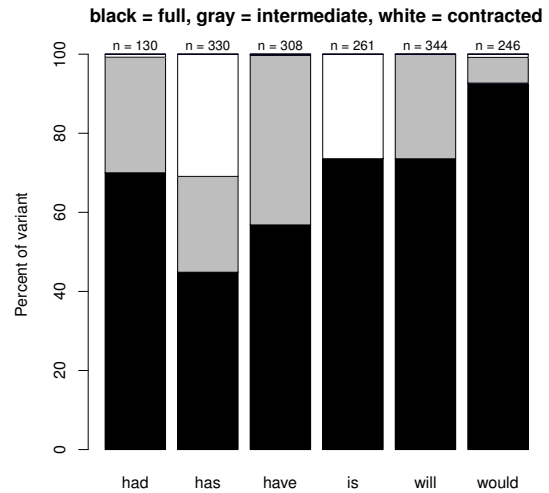**black = full, gray = intermediate, white = contracted**

Figure 2: After noun phrase subjects.

The plots above reveal that intermediate forms are in fact well-represented in natural speech data, despite the little attention that has been given them in the literature on contraction. An accurate model of contraction, then, needs to account for their existence. However, doing so is not trivial. Is contraction in fact a *tripartite* alternation underlyingly? That is, is each of the three surface forms for a given auxiliary a distinct allomorph? Or can we instead explain away intermediate forms as the output of processes unrelated to the alternation of full and contracted allomorphs, thus preserving the bipartite alternation put forth in earlier literature?

I turn to these questions in Sections 4.1–4.4. I will argue that the latter analysis is correct: specifically, that auxiliary realization necessitates two stages of processes, one a morphosyntactic alternation between underlying full and contracted forms, and another a set of low-level phonetic or phonological rules that operate on the output of that morphosyntactic alternation. It is the combination of these two stages of processes that generates intermediate forms. Localizing variation in two stages can thus account for the attested surface patterns.

4.1   Distribution of Intermediate Forms

First, a close inspection of Figures 1 and 2 reveals that not every host/auxiliary combination permits intermediate forms to surface. Descriptively, intermediate forms are found:

- As forms of *had* (and, to a much lesser extent, *has* and *will*) after vowel-final pronouns.
- As forms of all auxiliaries (but *is*, which has no intermediate form) after noun phrases.

These environments can be sorted into two categories, as follows:

(11)   Intermediate forms that surface where contracted forms **are also** possible:

   a. Pronoun + *had*: intermediate *he* [əd] surfaces alongside contracted *he'd* [hid]

   b. Noun phrase + *has*: intermediate *Sue* [əz] surfaces alongside contracted *Sue's* [suz]

(12)   Intermediate forms that surface where contracted forms **are not** possible. . .

   . . . for reasons of phonotactics:

   a. Consonant-final noun phrase + *had/will/would*: intermediate *Pat* [əd]/[əl] has no contracted counterpart *Pat'd*/*Pat'll* *[pætd]/*[pætl]

   b. Consonant-final noun phrase + *have*: intermediate *children* [əv] has no contracted counterpart *children've* *[ˈtʃɪldrɪnv]

   . . . for reasons other than phonotactics:

   c. Vowel-final noun phrase + *had/would/will*: intermediate *Sue* [əd]/[əl] has no contracted counterpart *Sue'd*/*Sue'll* *[sud]/*[sul]

   d. Vowel-final noun phrase + *have*: intermediate *three* [əv] has no contracted counterpart *three've* *[θriv]

In Sections 4.2–4.3 I will argue that intermediate forms in (11) are of a different source than intermediate forms in (12). Specifically, those in (12) come from the insertion of a contracted allomorph that fails to syllabify with its host, while those in (11) are not contracted allomorphs at all: they are full forms that have lost their initial consonant.

4.2   Intermediate Forms as Underlyingly Contracted Allomorphs

One crucial point to recognize is that just because an auxiliary does not surface in its contracted form — that is, does not surface in a single syllable with its host — does not mean that it was underlyingly a full form. Compelling evidence for this comes from the behavior of the auxiliary *will* after the pronoun *it*. As shown in Figure 1, *will* contracts at a high rate (91%) after vowel-final personal pronouns. After the pronoun *it*, on the other hand (not included in Figure 1), contracted forms are not found (unsurprisingly, given English phonotactics), but intermediate forms are found instead, at a comparably high rate (79%). This near-complementary distribution is easily explained if we allow that intermediate forms may derive from underlyingly contracted forms that undergo a repair process of epenthesis of ə (henceforth, ə-Insertion) when they cannot syllabify with their host. (See Chomsky and Halle 1968 for a similar process which they propose to account for alternations such as *carpentry*∼*carpenter*, *twinkling*∼*twinkle*.) The analysis presented in Section 2.2 of "contraction" as allomorph insertion followed by syllabification thus can account for intermediate forms in environments of phonotactic unacceptability: contracted forms can still be inserted in such an environment, but they simply fail to syllabify.

I extend this failure of syllabification to other environments given in (12). Contracted forms of auxiliaries apart from *is* and *has* in fact fail to syllabify with *any* non-pronoun noun phrase host, regardless of phonotactic acceptability (the *[məˈrid], *[θriv], *[sul] of (7–10)). The alternative to full forms for these auxiliaries in such cases is, again,

intermediate forms. So once again, I analyze these as cases where a contracted allomorph is inserted, but fails to syllabify, necessitating a ə-Insertion repair process. The alternative to this analysis is restricting the insertion of the *'ve, 'll, 'd* allomorphs to only post-pronouns. I propose instead that these allomorphs can occur after both pronoun and noun phrase hosts. Rather, it is the syllabification process that fails for them in this environment, obscuring the contraction that has occurred underlyingly.

A mention must be made of *why* syllabification fails in these latter cases. *[məˈrid], *[θriv], and *[sul] are phonotactically acceptable but ungrammatical. I am currently investigating an analysis under which syllabification of non-*'s* contracted forms fails to occur over boundaries or brackets that surround a host of a given weight. By this metric, noun phrase subjects have boundaries that exceed a critical threshold and prevent syllabification; pronouns do not. Related to this is the finding, currently being explored, that the longer a subject, the less likely contraction is to occur after it, such that contracted forms of *is* and *has* are essentially nonexistent after noun phrases of greater than eight words (see MacKenzie 2011). At any rate, the upshot of all this is that just because host-auxiliary syllabification is impossible in a given environment does not mean we can rule out host-auxiliary *adjunction* having occurred. Adjunction may have occurred but the contracted form may simply be unable to surface, for either phonotactic or "heaviness" reasons.

4.3 Intermediate Forms as Underlyingly Full Allomorphs

The previous section argued that intermediate forms in environments in which syllabification of a contracted form is illicit could still be traced to an underlyingly contracted source: host-auxiliary adjunction has applied, but syllabification has failed. This was used to account for intermediate forms of the (12) type. But (11)-type intermediate forms surface as well: cases in which syllabification of contracted forms is acceptable, as indicated by the appearance of contracted forms alongside the intermediate ones.

Two explanations present themselves for (11)-type intermediate forms. The first is that they are of the same source as (12)-type forms: host-auxiliary adjunction has occurred but syllabification has failed and ə-Insertion has occurred as a repair. For this to be the case, syllabification must be variable: sometimes it can occur (hence *he'd, John's*); sometimes it fails (hence *he* [əd], *John* [əz]). Alternatively, these intermediate forms may not be instances of contracted forms at all. Instead, they may be cases in which a full allomorph has been inserted, but separate processes have turned it into an intermediate form that simply happens to be phonologically similar to those that are derived from syllabification failure and ə-Insertion.

I propose that the second analysis is the one that is correct. That is, syllabification is not optional: it occurs whenever possible. Any case of an intermediate form in a (11)-type environment is an instance of a full form having been inserted and turned into an intermediate form. This analysis will turn out to explain an important asymmetry seen in Figure 1 regarding the patterning of *h*-initial and *w*-initial auxiliaries.

What processes could generate intermediate forms from full forms? We need one

process to reduce a full form's vowel to schwa, and another to delete a full form's initial consonant. Both such processes are in fact independently attested in English, and both are presented in Kaisse's (1985) discussion of fast speech rules. The first – Kaisse's Vowel Reduction — affects unstressed syllables and applies commonly to function words; the second is attested in Kaisse's process of *h*-Deletion, which deletes *h* when it begins unstressed syllables, applying, again, to pronouns and function words (e.g. *he, him*) in conversational speech. So, we now have a separate, and entirely different, source of intermediate forms: the output of low-level phonological and phonetic processes having applied on full forms.

Crucially, I follow Kaisse (1985) in assuming that there is no process that deletes initial *w* in a fashion analogous to *h*-Deletion. Intermediate forms of the *w*-initial auxiliaries *will* and *would* must have their source only in ə-Insertion on contracted forms: they are (12)-type only. Where the full form of these auxiliaries is inserted, I assume, it surfaces as-is. This provides a simple explanation for the conspicuous lack of intermediate forms of *would* after vowel-final pronouns in Figure 1: since this is an environment in which syllabification is acceptable, contracted forms surface as-is; since there is no *w*-deletion, full forms likewise surface as-is. Intermediate forms of *w*-initial auxiliaries thus cannot be derived after vowel-final pronouns, and are thus expected not to surface.[2] An analysis in which (11)-type intermediate forms are generated from full forms, then, neatly explains the observed asymmetry in distribution of intermediate forms between *h*-initial and *w*-initial auxiliaries, in a way that a "variable syllabification" account cannot easily do.

4.4   Summary

The derivations of the two types of intermediate forms laid out in Sections 4.1–4.3 are schematized in Table 1, with one example of each. In the (11)-type example, the host is *he*, the auxiliary is *had*, and variable adjunction does not apply. In the (12)-type example, the host is *Sue*, the auxiliary is again *had*, and variable adjunction does apply. Note that this is not intended to represent the only output possibilities for these particular inputs but rather to demonstrate how intermediate forms of the two different types would be generated. Processes that are variable are preceded by ∼.

	(11)-type D⌢Aux e.g. *he had*	(12)-type N⌢Aux e.g. *Sue had*
∼Adjunction	doesn't apply	[N[Aux]]
Vocabulary Insertion	he⌢hæd	[Sue[d]]
Syllabification	N/A	fails for "heaviness" reasons
ə-Insertion	N/A	Sue [əd]
∼Vowel Reduction	he [həd]	N/A
∼*h*-Deletion	he [əd]	N/A
Output	he [əd]	Sue [əd]

Table 1. Derivation of two types of intermediate forms.

[2]There are, contrary to this prediction, five intermediate forms of *will* after vowel-final pronouns in the corpus data. My working assumption is that these are due to diphthongization of the vowel of the pronoun when followed by *l*, and that the observed [ə] is a vocalic offglide rather than a schwa inserted for syllabification, but this awaits further investigation.

To summarize, then, the tripartite surface realization (full – intermediate – contracted) of the auxiliaries under study can be reduced to a bipartite distinction (full – contracted) underlyingly. Intermediate forms are simply full or contracted forms that have been subject to separate, parochial processes. Crucially, then, this analysis proposes that even where a contracted form is illicit on the surface (e.g. *[mə'rid], *[θriv], *[sul]), it may still have been inserted underlyingly. Its failure to surface is the result of what happens after the processes of adjunction and allomorph insertion have occurred.

4.5   Extensions of the Model

Having now proposed these two levels of variable processes — one level of morphosyntactic adjunction followed by one level of phonetic and phonological processes — the next step for a thorough account of auxiliary realization is to determine what the variable factors conditioning the application of those processes must be. Are the patterns found in natural speech a result of conditioning factors on the morphosyntactic rule of adjunction, on the process of *h*-Deletion on full forms, on both? This is obviously a much larger task than can be addressed in this paper, and one which requires data from many more independent variables than are examined here, but I lay out the preliminary issues below.

One pattern evident in Figure 1 but not yet touched on here is that auxiliaries pattern in two groups after pronouns: one group, consisting of *has, have, is,* and *will*, shows ≥90% contracted forms, while the other, consisting of *had* and *would*, shows ≤50% contracted forms. This finding also emerges, though it is not discussed, in McElhinny's (1993) study. Why should contraction show this dramatic effect? A convincing explanation turns out to be difficult to pin down. For example:

- **Word frequency.** Are lower-contracting auxiliaries less frequent? Frequency could successfully account for the high contraction rate of *is*, which is the highest contractor of the group (with 99% contracted forms after pronouns) and also the most frequent (with 62,095 occurrences in Switchboard), as well as the low contraction rate of *had*, which is the lowest contractor in the group (with 34% contracted forms after pronouns) and also the least frequent (with 1411 occurrences in Switchboard[3]). But this analysis fails to account for the patterning of *will* and *would*, which have contraction rates of 91% and 50% but Switchboard frequencies of 6,003 and 9,783, respectively.
- **Phonology.** Both low-contracting auxiliaries end in /d/. Is there something phonological about /d/ that disfavors contraction processes?
- **Past tense.** Both *had* and *would* bear a [+past] feature. Is the process of host-auxiliary adjunction disfavored when it would apply to an auxiliary that bears this feature?

The second two options are difficult to distinguish between. For now, I have chosen

---

[3]Both full and contracted forms of each auxiliary are included in these counts. Because Switchboard is not part-of-speech-tagged, there are cases in which a full or contracted auxiliary is indistinguishable from another, unrelated form (e.g. *'s* 'is' = *'s* [genitive marker]; *have* [auxiliary] = *have* [main verb]). In such cases, frequencies were approximated by pulling out 100 tokens at random of the ambiguous item, counting the proportion of relevant to irrelevant hits in that set of 100, and then using that ratio to estimate the number of relevant hits in the entire corpus.

to call the distinction one between "past" and "non-past" auxiliaries, but this is a topic for future work. Crucially, a [±past] condition is one that could be incorporated into the structural description of the morphosyntactic adjunction rule as it is currently written. An example of how this would be accomplished is roughly sketched in (13), in which the host-auxiliary adjunction process is probabilistically conditioned by the tense of the auxiliary.

(13) $X^\frown T \rightarrow [X[T]]$ $p_{[-past]} > p_{[+past]}$

While this conditioning factor is clearly in effect after pronoun subjects, the prevalence of intermediate forms after noun phrase subjects (Figure 2) obscures any evidence of the pattern in that environment. Attributing those intermediate forms to their underlying source will reveal whether or not the [±past] condition is in evidence after noun phrases as well: in other words, whether host-auxiliary adjunction is identically conditioned regardless of host, or whether there is instead an interaction of the past/non-past constraint with host type. Because the latter case would allow for the potential of there being *two* distinct adjunction processes, one per environment, identifying the extent of this factor's influence has implications for how the contraction alternation is represented in the grammar.

Intermediate forms can be recategorized as their underlying forms as follows:

- Intermediate forms of post-noun phrase *has*, being in the (11)-type group, are underlyingly full.
- Intermediate forms of post-noun phrase *will* and *would*, being in the (12)-type group, are underlyingly contracted.
- Intermediate forms of post-noun phrase *had* and *have* are of the (12) type as well. However, because the full forms of these auxiliaries begin with *h*, any full form of post-noun phrase *had* or *have* could variably undergo *h*-Deletion to surface as intermediate. Intermediate forms of *had* and *have* are thus ambiguous between the two possible sources.

In the unambiguous cases, this reclassification begins to uncover an underlying past/non-past split, exactly as seen after pronouns. Specifically, given reclassification of their intermediate forms as detailed above, *has, is,* and *will* all show underlying contraction rates of around 30%, while *would* shows one that is much lower — closer to 7%. This is very similar to the pattern that we saw after pronouns.

Where intermediate forms are ambiguously derived from the two possible sources, it is less immediately obvious how to determine their provenance. One method is to apply the rates of adjunction and *h*-Deletion shown in the unambiguous cases to the ambiguous cases, to produce a model full/contracted distribution. If this model does not differ significantly from what is observed, the hypothesis that the ambiguous auxiliaries undergo adjunction at the same rate as the unambiguous ones can be upheld.

This was undertaken with post-noun phrase *had* and *have*. The rate of adjunction used was 26%, that shown by the contracted forms of post-noun phrase *is* and *will*. The rate

of *h*-Deletion used was 35%, that shown by the intermediate forms of post-noun phrase *has*. These rates, when applied to the total number of forms of *had*, overestimate the number of intermediate forms that should surface: the model full/intermediate breakdown they predict differs significantly from that which is observed ($\chi^2 = 25.9322$, $p < .001$). This means that either one or the other rate is too high. Since *h*-Deletion is a low-level, fast-speech process that is not likely to show auxiliary-specific behavior, it is more likely that it is the rate of adjunction that is too high: in other words, that *had* undergoes adjunction to its host at a lower rate after noun phrases than do *has, is,* or *will*. This is the same pattern that was in evidence after pronouns: another indication that the past/non-past split may be in effect after noun phrases as well.

But applying the same rates to the intermediate forms of *have* gives the same result: the other auxiliaries' rates of host-auxiliary adjucntion and *h*-Deletion overestimate the number of intermediate forms of *have* relative to what is actually displayed in the data ($\chi^2 = 9.4819$, $p = .002$). Once again, either the adjunction rate or the *h*-Deletion rate shown by the other post-noun phrase non-past auxiliaries is too high. A step for future work is to examine why this might be, and to determine whether this means only that a past/non-past split after noun phrases must be rejected. There could be, for instance, some acoustic reason for *h*-Deletion to be lower for post-noun phrase *have*, which follows a final /s/ more often than do other post-noun phrase auxiliaries.

The main take-home point of this section is that much of the data shows an apparent past/non-past split, one that needs to be incorporated in any model of contraction processes. Given the analysis of contraction proposed here, in which contracted forms are contingent on morphosyntactic adjunction having applied, this means that such a condition would need to be represented on the adjunction rule itself. Future work, which will nail down the precise nature of what has been called here the "past/non-past" effect, will determine whether this conditioning factor is one that lends itself to a single adjunction process regardless of subject, or whether there is evidence for host-specific rules of adjunction.

5.   Conclusion

This paper has reported the results of a corpus study of English auxiliary contraction. Though much literature has been devoted to analyzing the mechanisms governing auxiliary realization, past analyses have crucially been based only on judgment data. As a result, any patterns displayed in natural speech have been overlooked, despite the potential implications they may have for the locus of contraction processes in the grammar.

One such pattern documented here is the prevalence of "intermediate" forms of auxiliaries, which do not fit neatly into the binary split between full and contracted allomorphs proposed in previous literature. I have argued here that these intermediate forms can be accounted for under a model of contraction that has two stages: one stage of variable morphosyntactic adjunction, and another stage of low-level processes that operate on the full and contracted forms that have been output by that initial stage.

A second pattern is a split in contraction rate by which, after pronouns, [-past] auxil-

iaries display contracted forms at a much higher rate than [+past] auxiliaries. I have shown that judicious treatment of post-noun phrase intermediate forms nearly replicates this split after noun phrase subjects as well, with the exception of forms of *have*. Future work will focus on how exactly this condition is to be incorporated into a model of contraction, and whether there is reason to propose distinct post-pronoun and post-noun phrase rules, or whether contraction processes are instead identically conditioned irrespective of subject.

References

Anderson, Stephen R. 2008. English reduced auxiliaries really are simple clitics. *Lingue e Linguaggio* 7:169–186.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. Cambridge, MA: The MIT Press.

Godfrey, John J., Edward C. Holliman, and Jane McDaniel. 1992. SWITCHBOARD: Telephone speech corpus for research and development. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Volume 1*, 517–520.

Inkelas, Sharon, and Draga Zec. 1993. Auxiliary reduction without empty categories: A prosodic account. In *Working Papers of the Cornell Phonetics Laboratory 8*, ed. Corinne Moore and Ann Bradlow, 205–253.

Kaisse, Ellen M. 1983. The syntax of auxiliary reduction in English. *Language* 59:93–122.

Kaisse, Ellen M. 1985. *Connected speech: The interaction of syntax and phonology*. New York: Academic Press.

King, Harold V. 1970. On blocking the rules for contraction in English. *Linguistic Inquiry* 1:134–136.

Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45:715–762.

MacKenzie, Laurel. 2011. Quantitative data as a clue to auxiliary contraction processes. Paper presented at the 2011 annual meeting of the Linguistic Society of America.

McElhinny, Bonnie S. 1993. Copula and auxiliary contraction in the speech of White Americans. *American Speech* 68:371–399.

Sells, Peter. 1983. Juncture and the phonology of auxiliary reduction in English. In *University of Massachusetts Occasional Papers in Linguistics 8*, ed. Toni Borowsky and Daniel Finer, 76–105.

Wilder, Chris. 1997. English finite auxiliaries in syntax and phonology. In *Clitics, pronouns and movement*, ed. James R. Black and Virginia Motapanyane, 321–362. Philadelphia: John Benjamins Publishing Co.

Zwicky, Arnold M. 1970. Auxiliary reduction in English. *Linguistic Inquiry* 1:323–336.

Department of Linguistics
619 Williams Hall
University of Pennsylvania
Philadelphia, PA 19104-6305

laurel@ling.upenn.edu