Locating linguistic variation: A case study of English auxiliary contraction

Laurel MacKenzie, University of Pennsylvania

This paper argues that quantitative data can be used to identify the locus of variation of complex morphosyntactic phenomena. We take as a case study English auxiliary contraction, a process which has been analyzed from many perspectives, ranging from phonological to syntactic. Despite the large body of theoretical work on contraction, however, the fact that this phenomenon is variable (e.g. *He had ∼ he'd been there all day*) has received little notice outside of the sociolinguistic literature. This oversight is unfortunate, as the factors that condition this variation have ramifications for how the process must be situated in the grammar. In this paper, we show that although there are a number of places this variation could in principle be located, the quantitative data argue for one analysis in particular. Specifically, we argue for a variable rule of the syntax that affixes an auxiliary to its host, feeding the insertion of a clitic form.

Early theoretical studies treated contracted forms as the outcome of phonological rules (Zwicky, 1970; Selkirk, 1984); more recent work (Kaisse, 1985; Inkelas and Zec, 1993; Close, 2004; Anderson, 2008) has abandoned this route to treat contracted auxiliaries as suppletive clitic allomorphs whose insertion is governed by syntactic or prosodic constraints. In no theoretical study, however, are data from large bodies of natural speech brought to bear on the analysis of this alternation. In the sociolinguistic literature, by contrast, empirical data have been used in many studies (beginning with Labov (1969)) to define a number of variable constraints on contraction, but the analysis of this variable has not moved past characterizing auxiliary contraction as a phonological rule.

We integrate these two approaches to contraction in our study, based on an examination of 4400 total tokens of the auxiliaries *had, has, have, is, will,* and *would* collected from the Switchboard corpus (Godfrey et al., 1992). Tokens were coded as full, as in (1a), intermediate, as in (1b), or contracted, as in (1c), as well as for linguistic features of the subject and for demographic characteristics of the speaker. The major point of comparison focused on in this study will be the effects of pronoun versus full noun phrase subjects.

The data, presented in Figures 1 and 2, reveal several important findings. First, as reflected in the coding scheme, the bipartite distinction between full and contracted traditionally assumed in the literature oversimplifies the case. Intermediate forms are well-represented—even in environments where fully contracted forms are also acceptable, such as *had* after pronoun subjects—and must be accounted for. We propose that these intermediate forms may come from two sources. They may be due to a low-level process of *h*-Deletion (Kaisse, 1985) that operates on the full forms of unstressed *h*-initial auxiliaries whose vowels have been reduced to schwa (2). Or, they may derive from underlyingly contracted forms that are unable to syllabify with their host and must surface with an epenthetic schwa (3). Attributing these ambiguous forms to one or the other process has different implications for the data, as discussed below.

Figure 1 also reveals the novel finding that after pronouns, auxiliaries do not all contract at equal rates. Specifically, the non-past auxiliaries *has, have, is,* and *will* are alike in showing high rates of contraction in this environment, while the past-tense auxiliaries *had* and *would* contract at a much lower rate. We propose, in the interest of parsimony, that this distinction between past and non-past auxiliaries holds after noun phrase subjects as well, though it is obscured there by the high rate of ambiguous intermediate forms. If we assume that most ambiguous forms of non-past auxiliaries in this environment are underlyingly inserted clitics (3), while most ambiguous forms of *had* are underlyingly full forms that have been *h*-Deleted (2), the finding that non-past auxiliaries contract at a higher rate than past-tense auxiliaries is accounted for.

We propose that these results lend themselves to a two-stage process that incorporates both syntactic and phonological rules. We follow Kaisse (1985) in positing a syntactic rule of affixa-

tion under adjacency that adjoins an auxiliary to its host, feeding the insertion of a contracted allomorph. Given our results, we add that this affixation rule is variably conditioned by both auxiliary tense and host type (NP or pronoun). Then, h-Deletion and ə-Insertion apply in the phonology, operating on full and contracted forms, respectively.

We conclude the paper with a discussion of further questions specific to contraction: for instance, how specifically to characterize the difference between non-past and past auxiliaries (as *had* and *would* do not behave identically and thus are not necessarily forming a coherent class). More generally, we also discuss what variation can tell us about the grammatical architecture, taking the results of this study as evidence for a hierarchical grammatical model.

(1)   a. Full: He [hæd]/[həd] been there all day. (initial consonant, audible vowel)

   b. Intermediate: He [əd] been there all day. (no initial consonant but audible vowel)

   c. Contracted: He[d] been there all day. (no initial consonant, no vowel)

(2)   Full form + *h*-Deletion:
      /hæd/ → [hæ̆d] → [hə̆d] → [ə̆d]

(3)   Contracted form + ə-Insertion:
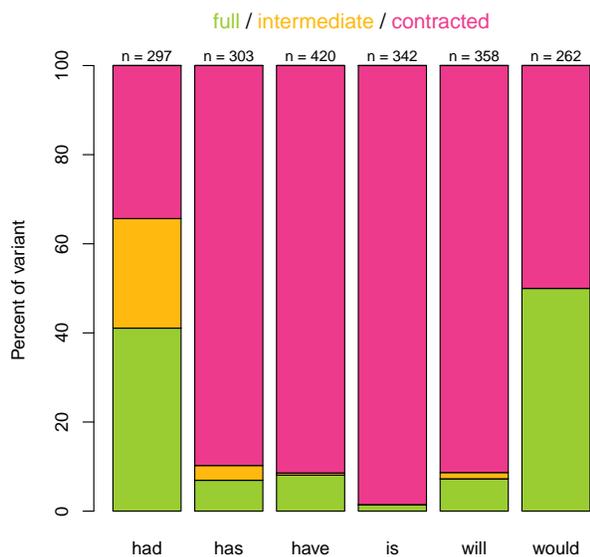   two /d/ → two [əd] 'two had'

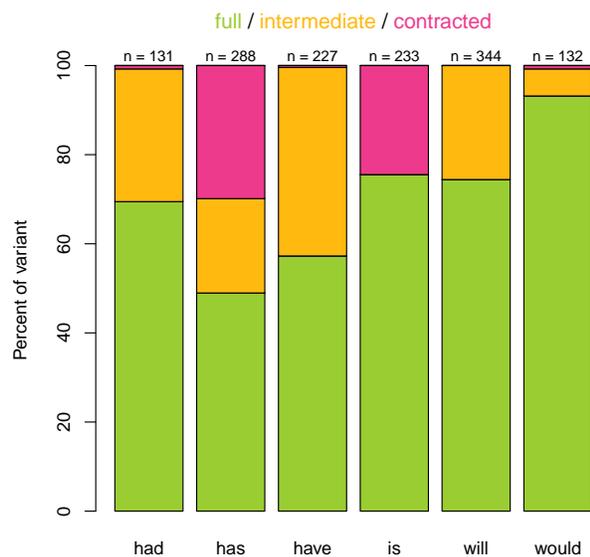
Figure 1: After pronoun subjects.

Figure 2: After noun phrase subjects.

References. Anderson, Stephen R. 2008. English reduced auxiliaries really are simple clitics. *Lingue e Linguaggio* 7:169–186. ● Close, Joanne. 2004. English Auxiliaries: A Syntactic Study of Contraction and Variation. Doctoral Dissertation, University of York. ● 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. 1985. Connected Speech: The Interaction of Syntax and Phonology. New York: Academic Press. ● Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45:715–762. ● Selkirk, Elisabeth O. 1984. *Phonology and Syntax: The Relation Between Sound and Structure.* Cambridge, MA: The MIT Press. ● Zwicky, Arnold M. 1970. Auxiliary reduction in English. *Linguistic Inquiry* 1:323–336.