English auxiliary contraction as a two-stage process: Evidence from corpus data
keywords: contraction, variation, corpus study

Overview. The variable contraction of English auxiliaries (e.g. John is ~ John’s coming) has been well studied (Zwicky, 1970; Kaisse, 1983; Anderson, 2008, a.o.). However, outside of the sociolinguistic domain, analyses of this phenomenon have been based on judgments, not quantitative data. In this paper, we show that considering the quantitative distributions of contracted forms results in a more nuanced picture of contraction, one that has implications for where the process must be situated in the grammar. Specifically, we present data from a corpus study of contraction that support a two-stage analysis of the phenomenon: an allomorphic alternation governs the insertion of contracted and non-contracted forms; then, subsequent phonetic and phonological processes operate on the inserted allomorph.

Methodology. We examined 1075 tokens of post-noun phrase has, have, is, and will collected from the Switchboard corpus (Godfrey et al., 1992). Tokens were coded as full (1a), intermediate (1b), or contracted (1c), as well as for linguistic features of the subject. Henceforth, we will use the terms “full,” “intermediate,” and “contracted” to refer to forms of the (1a), (1b), and (1c) types, respectively.

(1) a. Full: John [hæz]/[hæz] been there all day. (initial consonant, audible vowel)
  b. Intermediate: John [z] been there all day. (no initial consonant but audible vowel)
  c. Contracted: John[z] been there all day. (no initial consonant, no vowel)

Findings and analysis. Our data confirm earlier findings that contracted forms of auxiliaries other than is and has are impossible after noun phrases: *Sue’ll [sul] ‘Sue will’, *three’ve [thriv] ‘three have’. Instead, intermediate forms of will and have, and of has as well, surface after noun phrase subjects (Table 1).

<table>
<thead>
<tr>
<th></th>
<th>has</th>
<th>have</th>
<th>is</th>
<th>will</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent full forms</td>
<td>49</td>
<td>57</td>
<td>75</td>
<td>74</td>
</tr>
<tr>
<td>Percent intermediate forms</td>
<td>21</td>
<td>42</td>
<td>N/A</td>
<td>17</td>
</tr>
<tr>
<td>Percent contracted forms</td>
<td>30</td>
<td>1</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>Number of tokens</td>
<td>287</td>
<td>226</td>
<td>230</td>
<td>332</td>
</tr>
</tbody>
</table>

Table 1. Distribution of forms.¹Is has no intermediate form distinct from its full form.

Under previous analyses of contraction as a bipartite alternation between full and contracted allomorphs (e.g., Kaisse, 1983), the source of these intermediate forms is not immediately apparent. But a thorough account of the contraction process requires that they be explained. We account for the existence of intermediate forms by positing two processes that act on the output of the allomorphic alternation between full and contracted. One process is fast-speech /h/-Deletion, which creates intermediate forms from unstressed full forms of has and have (2). This process is independently attested in English as applying to function words such as he and him (Zwicky, 1970). The other is [ə]-Epenthesis, which inserts [ə] before a contracted allomorph that fails to syllabify with its host. [ə]-Epenthesis allows contracted forms to be syllabified where they would otherwise form a phonotactically illicit cluster with their host (3a); it likewise syllabifies contracted forms of will and have, which cannot attach to a noun phrase despite phonotactic acceptability (3b).

(2) Full form + /h/-Deletion: /hæz/ → [hæz] → [həz] → [əz]

(3) a. Contracted form + [ə]-Epenthesis for phonotactic reasons: Pat /l/ → Pat [əl] ‘Pat will’
   b. Contracted form + [ə]-Epenthesis despite phonotactic acceptability: Sue /l/ → Sue [əl] ‘Sue will’
Extensions. The upshot of this analysis is that the underlying full/contracted distinction is, for all auxiliaries except is, obscured on the surface by /h/-Deletion and [a]-Epenthesis. In the case of have, this results in intermediate forms whose provenance is ambiguous: both /h/-Deletion and [a]-Epenthesis create surface-identical intermediate forms of have. But intermediate forms of will and has can be unambiguously traced to a distinct underlying source: intermediate forms of will are strictly underlying contracted forms (no process of /w/-deletion exists to generate intermediate forms of will from full; see Kaisse, 1985), while intermediate forms of has are strictly underlying full forms (as [a]-Epenthesis is never necessary to syllabify /z/ with its host).

This means that any conditions governing the underlying alternation between full and contracted allomorphs will manifest themselves in the surface realizations of full and contracted particular to an auxiliary. This is borne out in the effect of DP length on auxiliary realization (Figure 1): contracted forms of is/has and intermediate forms of will are all disfavored with longer DPs, implying that, at an earlier stage in the derivation, they shared a source. Intermediate forms of have, however, show no such effect—unsurprising under the present analysis, which identifies these surface forms as a hybrid class: they come from both full and contracted underlyingly.

Figure 1. Effect of DP length on auxiliary realization. Longer DPs disfavor underlying contractions.

References