I can't understand - The perception of native and non-native can and can’t by native and non-native listeners of English

Margot van Mulken¹, Huib Kouwenhoven¹, Mirjam Ernestus¹²
¹Centre for Language Studies, Radboud University, Nijmegen
²Max Planck Institute for Psycholinguistics
Contact email: m.vanmulken@let.ru.nl

The difference between an affirmative and a negative English statement is realized by only one sound at the segmental level, as in I can/can’t go. In addition, there may be subsegmental cues in the surrounding segments, depending on the speaker’s native language background. This paper investigates the comprehension of /t/-reduction in American and Spanish English can’t by listeners with different language backgrounds, in order to increase our understanding of native and non-native speech perception.

Previous research suggests that /t/-reduction is not only common in native (American) English, but also occurs in non-native (L2) English speech. On the one hand, there are languages where absence of /t/ may result in word pronunciation variants that better match the phonotactic constraints of the speakers’ native languages (L1). For instance, in Spanish and Chinese, consonant clusters are phonotactically illegal, and word-final /t/ in word-final consonant clusters are therefore expected to be absent in English produced by Spanish and Chinese speakers as well. On the other hand, there are languages (e.g. Dutch) that reduce word-final /t/ as English does, and these speakers are likely to transfer this reduction phenomenon to their non-native English. The question arises how listeners with different language backgrounds cope with absent /t/ in native and non-native English. Do listeners who are used to /t/-deletion also easily reconstruct absent /t/? And how sensitive are listeners of different language backgrounds to subsegmental cues to absent /t/?

In an experiment, we investigated the comprehension of can, unreduced can’t and reduced can’t in American L1 English and Spanish L2 English speech. The tokens were all taken from spontaneous speech corpora: the American L1 English tokens from the Buckeye Speech Corpus (Pitt et al., 2007) and the Spanish L2 English tokens from the Nijmegen Corpus of Spanish English (NCSE, Kouwenhoven, Ernestus & van Mulken, in press). Tokens of can’t were classified as “reduced” if an automatic speech recognizer transcribed these tokens without [t]. The original orthographic transcriptions of these reduced tokens as tokens of can’t were verified by eight native speakers of English based on the orthographic context. In total, the experiment contained 218 tokens of can (95 from the NCSE), 147 tokens of unreduced can’t (48 from the NCSE) and 93 tokens of reduced can’t (43 from the NCSE).

In both corpora, can and can’t were produced with the same full vowel. Vowel quality could therefore hardly function as a cue for word identification. Durational analysis revealed that, in the Spanish reduced can’t tokens, the remaining segments (corresponding to can) were as long as the segments in Spanish English tokens of can. The Spanish reduced can’t tokens were thus very similar to the can tokens both at the segmental and durational level. Listeners may therefore be likely to interpret these tokens as can, independent of their language backgrounds. In contrast, in the American English tokens of can’t, the remaining segments were as long as the segments in unreduced can’t and shorter than the segments in can. The American English reduced can’t tokens were thus similar to can at the segmental level but similar to unreduced can’t at the durational level. As a consequence, listeners may differ in how they interpret American English reduced can’t, depending, among other things, on whether they are sensitive to fine durational cues.

The can / can’t tokens were presented together with the nouns that preceded them and with the verbs that followed them in the corpora. Participants indicated whether they had perceived can or can’t. We tested one native and three non-native listener groups. The non-native listeners were 1) Spanish listeners of English, who share the L1 background of the L2 speakers; 2) Mandarin-Chinese listeners of English, whose L1 belongs to a different language family entirely, but whose L1, like Spanish, does not have consonant clusters nor /t/-deletion; and 3) Dutch listeners of English, whose L1 belongs to the same language family and shows the same /t/-reduction pattern as English. A total
of 127 participants took part in our experiment. The Spanish were tested in Spain, the Dutch in the Netherlands, and the Chinese and American English listeners in North America.

Figure 1 shows the percentages of errors made by the four listener groups on the three stimulus types from each corpus. It clearly shows that the native listeners had little problems classifying the American English tokens, including the reduced tokens of can't. The Dutch show similar results, although their error percentages are generally a bit higher. These results suggest that these listeners relied on the subsegmental details that are present in American English reduced can't.

The reduced can't tokens produced by the Spaniards were more difficult for the American English and the Dutch listeners than the American English reduced can't tokens. This supports our conclusion from the durational analysis that the Spanish reduced can't tokens are more ambiguous than the American English ones at the subsegmental level and our hypothesis that Dutch and English listeners make use of the subsegmental cues present in the American English tokens. The Spaniards and the Chinese listeners, in contrast, showed very high error rates for American English reduced can't. These listeners apparently used the subsegmental cues in these reduced tokens to a much lesser extent. Nevertheless these listeners also classified reduced tokens of can't quite often as can't, rather than consistently as can. This shows that these listeners did take into account that /t/ may be absent, as also appears from the high error rates produced by the Spaniards for American English can.

Finally, the Spaniards did not understand their own speech better than the other listener groups. They generally pattern with the Chinese listeners, except for unreduced can't. The Chinese found it harder to identify /t/ and had a preference for can (which is also evidenced by their relatively low error rates for American English can).

In summary, the American English reduced tokens of can't were generally well understood by native listeners of American English and the related language Dutch. These listeners were able to make use of subsegmental traces left by reduced /t/. Native speakers of Spanish and Chinese, in contrast, used these cues to a lesser extent and often confused reduced can't with can. This was even the case for tokens produced by speakers with the same language background, probably because these tokens do not contain the relevant subsegmental cues. We conclude that one of the reasons why non-native listeners may have problems understanding reduced pronunciation variants is their inexperience with interpreting relevant subsegmental cues.

**Figure 1** Mean inaccuracies to the can and can’t stimuli; split by corpus (Nijmegen Corpus of Spanish English versus the Buckeye corpus), stimulus type and listener group.

**References**