On the historical origin of marginal contrasts: Canadian raising in Illinois
José I. Hualde, Tatiana Luchkina, Christopher D. Eager, Sarah Little
University of Illinois at Urbana-Champaign

Whereas it is probably true that instances of marginal or quasi-phonemic contrasts are to be found in every language [5],[7], [8],[14], it is also true that in any language phonemic analysis is for the most part straightforward. A question that arises, then, is how quasi-phonemic contrasts develop. Here we are concerned with contrasts between sounds that are partially in complementary distribution, where, furthermore, individual speakers within a single speech community differ in the categorization of specific lexical items and in the strength of their intuitions. These cases include, for instance, the diphthong-hiatus syllabification contrast in Castilian Spanish [1], [9], [15] and the /ai/ vs /ʌɪ/ contrast in some varieties of North American [2], [3], [4], [6], [10], [13], and Scottish English [14].

We will focus on the /ai/ vs /ʌɪ/ contrast in Chicagoland English, for which we have undertaken both production and perception experiments (we include also participants from Central Illinois for comparison). Kilbury [12] offers a description of his own Chicago idiolect, providing minimal and near-minimal pairs. Preliminary observation shows that speakers from Chicago, however, vary considerably in their awareness of the contrast and perhaps in its distribution and implementation. Participants were asked to read sentences containing 95 words with the target diphthong, including 15 words with a t-flap (e.g. writer) and 15 words with a d-flap (e.g. rider), as well as additional sentences constructed to test the effects of following context on three other vowels. As shown in Fig. 1 (for 15 Chicagoland participants), a following voiceless consonant and also a following t-flap conditions a higher vowel and a higher glide than a following voiced consonant. In particular, the nucleus of the diphthong has a higher F1 before d-flap than before t-flap (p < .001). Differences were also found between these two contexts in glide F2 (p < .002), but not in duration, although the difference in the duration of the diphthong preceding other voiced vs voiceless consonants is highly significant (p <.0001). That is, vowel quality differences between the two pre-flap contexts are maintained in spite of a lack of durational differentiation, showing that the vowel quality contrast has been phonologized. For comparison, an ANOVA on /ei/ in pre-flap contexts (e.g. trader/traitor) did not reveal any significant differences in formants or duration.

When we consider each speaker separately, it is apparent that for all of them voiceless consonants tend to condition higher realizations than voiced ones. Regarding the two pre-flap contexts, on the other hand, there is substantial variation among speakers, ranging from almost complete separation to extensive overlap. In comparison, Central Illinois speakers do not show any significant contrast between the two pre-flap contexts and also show extensive overlap before other voiced vs voiceless consonants.

In our perception study, 35 participants were presented with pairs of words on a computer screen and were asked to identify the word they heard by pressing a key. The two options given were always a potential minimal pair. Besides the Target items (e.g. writer/rider), Easy (e.g. write/ride) and Difficult (homophones, e.g. petal/pedal) pairs were used. The items were extracted from the production experiment and spliced onto a carrier phrase. Data from 4 speakers were used for
the stimuli. Reaction times were significantly higher for Target than for Easy or Difficult items. Accuracy for Target items was somewhat higher than chance: 61.8% (cf. Easy = 91.1%, Difficult = 48.1%). A mixed effects regression analysis shows that nucleus F1 and glide F2 predict the patterns of responses and that the relative frequency of each pair member also had an effect.

We conclude that the /aɪ/ vs /ʌɪ/ contrast is “quasi-phonemic” at the level of the speech community. In the context before an alveolar flap, speakers differ in the consistency with which a difference is implemented and possibly also in the lexical distribution of the diphthongs. This is also manifested in perception, where greater uncertainty about possible contrasts causes longer reaction times than both for other contrasts and for homophonic pairs, as well as only moderate success in word identification without a context.

In English, vowels are considerably shorter before voiceless consonants than in other contexts. For the diphthong /aɪ/, reduced duration produces a decrease in the phonetic distance between nucleus and glide. In some dialects, including Chicago, this difference in quality has been exaggerated, promoting a contrast based on the nature of the underlying consonants before flapped alveolar stops, where no quantity difference is found. The resulting surface opacity in the distribution favors the categorization of the two diphthongs as contrastive, allowing for further changes in their lexical distribution.

It has been argued that full-fledged phonemic splits require the existence of a prior stage with distinct allophones in complementary distribution [11]. Phonemic split takes place when the conditioning context is subsequently lost in sound change. Marginal contrasts of the type considered here arise by the same evolutionary mechanism, except that the conditioning environment is still present in the lexical items where the contrast arose. Consequently, the distribution of the semi-contrasting categories is relatively predictable.

![Figure 1. Mean F1 and F2 values of /ai/ nucleus and glide by following context (voiced consonant, voiceless consonant, d-flap, t-flap). Left, female (n=10); right, male (n=5).](image)
References


