

## ABSTRACT

Vowel Duration as a Cue to the Perceptual Separation of Cognate Sounds in American English\*

Lawrence Raphael<sup>+</sup>  
Haskins Laboratories, New Haven

Much research by linguists and phoneticians has been directed toward discovering a single physiological or acoustic basis which underlies the perceptual separation of cognate sounds. Experimental evidence indicates that neither of the two most commonly suggested bases, the voiced/voiceless opposition and the fortis/lenis opposition, has enough generality of distribution to be a unique explanation of cognate perception. There is, however, a substantial number of different acoustic and articulatory cues which can account for cognate perception in virtually all phonetic environments. These cues, some of which are thought of as reflexes of the voiced/voiceless or fortis/lenis oppositions, include stop release, aspiration, stop closure duration, friction duration, intraoral air pressure, preceding vowel duration, presence or absence of fundamental frequency, and timing relationships between articulation and glottal states. The last of these cues, currently expressed as voicing onset time (VOT) in the case of prevocalic stops, may be sufficiently extended in its generality to account for cognate perception in all phonetic environments. Or, there may be no way to generally account for the phenomenon.

The relationship between articulation and glottal states has received little attention for sounds in absolute final position. It is possible that voicing offset time, the mirror-image analogue of VOT, is a significant cue to cognate perception in this position. Spectrographic measurements of voicing offset time reliably separate cognate categories of stops and fricatives. Tape-cutting experiments, however, in which subjects heard a randomized series of real-speech words from which varying degrees of voicing during final /bdg/ closure was removed, revealed no significant changes in perception in the direction of /ptk/, even when the closure period was completely eliminated and a small portion of the preceding vowel transition was cut away.

Stop and fricative cognates are also reliably separated by spectrographic measurements of preceding vowel duration. Tape-cutting experiments on real speech, in which the vowels before final stops and fricatives were shortened, caused perception to change from /bdgvɔz/ to /ptkfθs/. Secondary perceptual reversals, for stimuli in which the vowel was virtually eliminated,

---

\*Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

<sup>+</sup>Also City University of New York.

indicated that voicing during final consonant closure does have minor cue value.

A variety of minimal and subminimal CVC(C) pairs was synthesized on the Haskins Laboratories' Pattern Playback. The vowel duration in each was varied over a range of values derived from those found in real speech. It was found that, regardless of the cues for voicing or voicelessness used in the synthesis of the final consonant or cluster, listeners perceived the final segments as voiceless when they were preceded by vowels of shorter duration and as voiced when they were preceded by vowels of longer duration. Discrimination tests revealed that perception across and within phoneme boundaries was continuous rather than categorical. It was also found that the cue of vowel duration is more effective before final stops and clusters than before final fricatives. The indication that voicing during final consonant closure does have minor cue value for cognate perception received further confirmation in the synthetic-speech experiment.

It is concluded that preceding vowel duration is both a sufficient and necessary cue to the perception of the voicing characteristic of word-final consonants and clusters: that is, word-final cognate sounds are perceptually separated and identified on the basis of the duration of the vowels which precede them. Such a finding runs counter to the traditional assumption that features which identify linguistic segments are produced within the articulatory period of those segments. Thus, traditionally, the expectation would have been that the perception of final voiced stops and clusters, for example, results from the listener's recognition of vocal pulsing during consonant closure. Instead, this study demonstrates that such a perception depends primarily upon the listener's recognition of a suitably long vowel preceding the closure for the final consonant or cluster.