

TECHNICAL CORRESPONDENCE

CROSS-VOWEL PHONOTACTIC CONSTRAINTS

Fudge (1969), Clements and Keyser (1983), and Davis (1984) have all pointed out a number of systematic phonotactic constraints holding between a prevocalic and a postvocalic consonant in English monosyllabic words. One of the strongest of these constraints, and one that has been observed by all three of the above mentioned researchers, is that there are no monosyllabic words of the form sCVC in which the same non-coronal (labial or velar) consonant flanks both sides of the vowel. Hence there are no English words like "spep" or "skik". Another constraint, noted by Davis (1984), is that there are no monosyllabic words of the form sNVN (where N can be any nasal consonant). Thus there are no words in English like "snam" or "sming". Here, I point out that these two constraints are in fact more general. Both these constraints are more general in that they are not just constraints on monosyllabic words or on single syllables, rather they are constraints on any sequence of sCVC (or sNVN) regardless where in the word (or, rather, morpheme) that sCVC sequence (or sNVN sequence) occurs. Also, the constraint on sCVC sequences is not just a constraint on identical consonants flanking both sides of the vowel, but on homorganic consonants (i.e., consonants having the same place of articulation) flanking both sides of the vowel.

Are the systematic constraints on sCVC sequences (in which the Cs are identical noncoronal consonants) and sNVN sequences (in which the N is any nasal consonant) which hold for English monosyllabic words actually general constraints on English syllables as is assumed by Davis (1984) (and also by Clements and Keyser 1983)? If, in fact, these are constraints on English syllables, one would expect to find words containing the sequence sCVCV (or sNVNV) since the postvocalic C (or N) would not be part of the initial syllable. So, for example, one might expect that there would be words like "skicky" or "spapoon" in which the postvocalic consonant is not part of the initial syllable, but there would not be words like "skick" or "spap" in which the postvocalic consonant is part of the initial syllable. If, on the other hand, the constraint on sCVC sequences (and sNVN sequences) is actually a constraint on a sequence of sounds, regardless of whether the sounds are all in the same syllable, then

possible words or sequences like "skicky" or "spapoon" would be non-occurring or at least extremely rare. A search was done on a computerized lexicon containing nearly 20,000 words from Webster's Pocket Dictionary to see if the sequences sCVC and sNVN occur in any polysyllabic words. The only word in this lexicon in which the sequence sCVC is found (where the Cs are identical non-coronal consonants) is the word "dyspepsia" where the sequence "spep" occurs. No other such words were found. Polysyllabic words having the sequence sCVC where the two Cs are not identical are much more common. A search through the 20,000-word lexicon gives us such words as "spaghetti", "scaffold", "scuba", "eskimo", and "episcopal". Thus it appears that the constraint on sCVC sequences is not really a constraint only holding within a syllable, but is a constraint on a sequence of sounds holding within a word. (Notice that in the exceptional word "dyspepsia" the sequence "spep" spans a morpheme boundary since the *s* is part of the prefix "dys-", which occurs in words like "dysfunction", "dysphasia", and "dystrophy".)

At first glance, the search through the lexicon of polysyllabic words containing the sequence sNVN suggests that the constraint on the sequence sNVN does not hold for polysyllabic words, unlike the constraint on sCVC sequences. The following 12 words containing the sequence sNVN were found: casement, congressman, dismantle, emplacement, fastening (with the orthographic *e* between the *t* and the *n* being deleted in pronunciation), marksman, placement, pronouncement, replacement, spokesman, statesman, and talisman. However, these words, like the exceptional "dyspepsia" mentioned above, are not monomorphemic; all of these words (with the possible exception of "talisman") involve morpheme boundaries between the *s* and the following nasal consonant. These data thus indicate that the constraint on sNVN sequences (as well as on sCVC sequences) are constraints on a sequence of sounds that hold within morphemes; they can be considered morpheme structure constraints, not word level or syllable structure constraints.

The constraint disallowing (monomorphemic) sCVC sequences (in which the Cs are identical non-coronal consonants) on further investigation turns out to be a more general constraint in that the two Cs do not have

Copyright 1989 by the Association for Computational Linguistics. Permission to copy without fee all or part of this material is granted provided that the copies are not made for direct commercial advantage and the *CL* reference and this copyright notice are included on the first page. To copy otherwise, or to republish, requires a fee and/or specific permission.

0362-613X/89/010109-110\$03.00

to be identical; rather, they cannot be articulated in the same place in the vocal tract. That is, there are virtually no monomorphemic words in English that have the sequence sCVC where the two Cs are either both labial or both velar. The only word in the 20,000-word lexicon that was found to violate this constraint is the word "skunk". The words "skag", "spam", and "spumoni" would also violate the constraint although they were not listed in the lexicon. That this constraint really does involve identical place of articulation is made evident when we consider the situation where the two Cs in an sCVC sequence are not homorganic. A search through the 20,000-word computerized lexicon revealed that no constraint whatsoever held when the two Cs were made at different locations in the vocal tract. For example, the sequence skV was followed by a labial consonant in 58 entries (e.g., "skip", "scuba"), an alveolar consonant in 151 entries (e.g., skit, skate), and a palato-alveolar consonant in 25 entries (e.g., scotch, sketch). The fact that there were virtually no words with a velar consonant following an skV sequence is of interest. Moreover, the sequence spV was followed by a velar consonant in 56 entries (e.g., spike, spook), an alveolar consonant in 196 entries (e.g., spit, speed), and a palato-alveolar consonant in 20 entries (e.g., speech, special); there were virtually no words where a labial consonant followed an spV sequence. Thus it is concluded that the constraint on sCVC sequences originally formulated by Clements and Keyser (1983) and Davis (1984) as a constraint on the occurrence of identical non-coronal consonants is in fact a more general constraint on consonants made in the same place of articulation.

Although the constraint against having homorganic (non-coronal) consonants flanking both sides of the vowel in a sCVC sequence seems to be a real constraint of English, it remains somewhat of a mystery why there should be such a constraint. The constraint crucially must include *s* since there is no constraint on English

CVC sequences where the two Cs are homorganic. A check through the 20,000-word computerized lexicon found 118 entries for words having (non-nasal) labial consonants flanking both sides of the vowel in a CVC sequence and 138 entries for words having a velar consonant flanking both sides of a vowel in a CVC sequence. Thus this constraint only involves an sCVC sequence and not any CVC sequence. I offer no explanation for why the presence of the *s* in an sCVC sequence essentially places a restriction on the postvocalic consonant. It is conjectured, though, that while the reason for such a constraint is a mystery, speakers of English make use of them for parsing words in continuous speech. For example, given the constraint on sCVC sequences discussed in this paper, phonetic sequences like [spalpleIn] (i.e., spy plane) and [ItsmaIn] (i.e., it's mine) can only be parsed as "spy plane" and "it's mine", respectively. They cannot be parsed as "spipe lane" and "it smine", respectively, nor could they be parsed as single words. It is quite possible that speakers of English can and do make use of such phonotactic constraints.

Stuart Davis
Speech Research Laboratory
Department of Psychology
Indiana University
Bloomington, IN 47405

ACKNOWLEDGMENTS

I wish to thank David Pisoni for his comments and Michael Dedina and Luis Hernandez for their technical assistance. This research was supported by an NIH Training Grant NS-07134-09 to Indiana University.

REFERENCES

- Clements, George N. and Keyser, Samuel J. 1983 *CV Phonology*. MIT Press, Cambridge, MA.
Davis, Stuart M. 1984 Some Implications of Onset-Coda Constraints for Syllable Phonology. *Chicago Linguistic Society* 20: 46-51.
Fudge, Erik C. 1969 Syllables. *Journal of Linguistics* 5: 253-287.