SPEECH ERROR EVIDENCE ON THE ROLE OF THE VOWEL IN SYLLABLE STRUCTURE

Erin Rusaw & Jennifer Cole

University of Illinois at Urbana-Champaign, USA
erusaw2@illinois.edu; jscole@illinois.edu

ABSTRACT

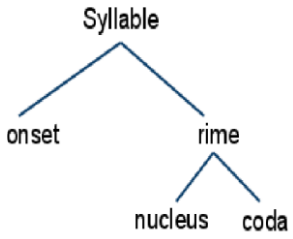
To explore the role of syllable structure in speech production, we conducted a study comparing C and V errors elicited through fast, paced repetition of nonsense CVC syllables. Subjects produced disproportionally more errors on consonants than on vowels, and more errors on onsets than on codas. There was no significant difference between the number of errors over CV and VC sequences. Also, most vowel errors were produced in conjunction with an error on one or both adjacent consonants, while most consonant errors were produced in isolation. These findings are at odds with the model of hierarchical syllable structure, especially in the disparity between singleton vowel and consonant errors. The observed pattern of speech errors is more consistent with the syllable model of Articulatory Phonology, where the vowel is not simply a sub-constituent of the rime, but plays a more central role in the coordination of consonantal elements with the vocalic nucleus of the syllable.

Keywords: syllable structure, speech errors, articulatory phonology, vowels

1. INTRODUCTION

Speech errors shed light on the nature of phonological encoding. Prior studies investigate phone-size segments and distinctive features as the basic units of phonological planning, and the organization of segments into syllables [2, 3, 4, 5, 8, 9]. In an early study of 179 naturally occurring speech errors, McKay [6] observes that syllable-initial C errors are more frequent than syllable-final C errors, which he attributes to syllable structure: final Cs are part of the rime constituent and errors are less likely to break constituents. He also observes that V errors are less common than C errors, which he explains in the same way. Although McKay's study forms the basis of much subsequent work on speech errors, his analysis of the C/V asymmetry is problematic. First, the hierarchical syllable model (Fig. 1) does not clearly predict an overall C/V asymmetry, and does not predict different likelihood of errors for coda consonants vs. vowels.

Figure 1: The hierarchical syllable model.



Second, McKay's findings are based on errors noticed in natural speech, which raises the question of whether vowel errors are truly less common, or whether vowel errors in conversational speech are simply less salient to listeners than consonant errors. In the current study we use a more controlled study of speech errors in lab speech to investigate whether there are in fact differences in the frequency of errors for different syllable positions, or between vowels and consonants overall. We then compare models of syllable structure to see which model(s) can best explain any differential error rates between syllable components.

2. METHODS

Syllables were constructed from a set of 6 onset Cs, 4 Vs and 4 coda Cs. Phrases were assembled from the syllables with four CVC syllables in each phrase with no repeated consonants or vowels (e.g. *vas pon dum fing*). 10 speakers of American English with no speech impairments viewed 45 distinct phrases presented orthographically, one at a time, on a computer screen. Subjects repeated each phrase six times along with a 200bpm metronome. Productions were audio recorded in a sound-dampened booth and later transcribed phone-by-phone by transcribers with experience in phonetic transcription. Transcribers were able to replay the data recordings as many times as necessary to ensure accurate transcription, and

were instructed to record a "?" if they were unsure which phone was being produced; all "?"s were discarded and not included in the data. Two transcribers each transcribed half of the data and a third transcribed 10% of the data assigned to each of the primary coders. There was a 95% agreement rate between the third transcriber and each of the two primary transcribers. The transcriptions were then compared with the list of stimuli (the intended productions), and any phone in a production which did not match the intended phone was marked as an error.

3. RESULTS

We found that subjects produced disproportionally more errors on consonants than on vowels ($c2(1, N = 1177) = 402.18$, $p < 0.0001$), and more errors on onsets than on codas ($c2(1, N=1066)=173.45$ $p<0.0001$). See Figures 2 and 3.

Figure 2: Speech errors on vowels vs. consonants.

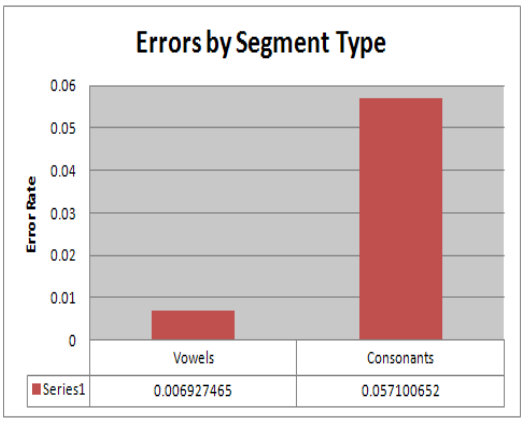
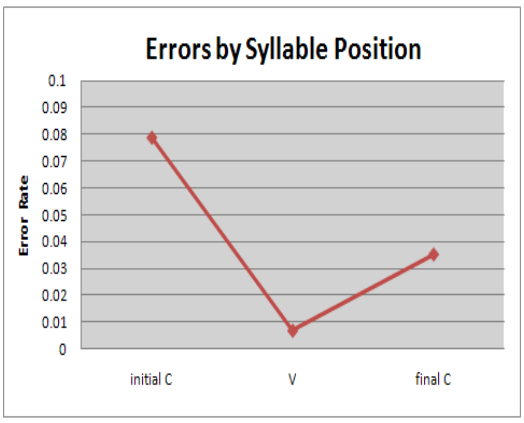


Figure 3: Speech errors on onsets vs. nuclei, vs. codas.

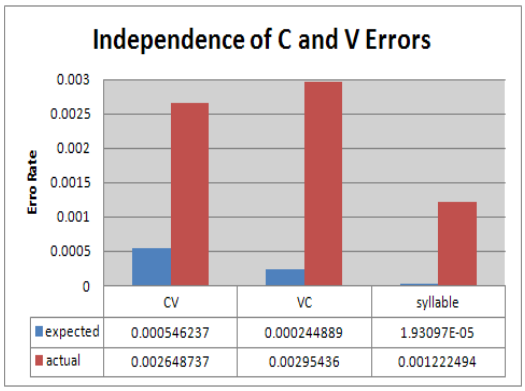


4. PDF DETAILS

Subjects also produced a significantly lower rate of errors on nuclei than codas ($c2(1, N=415) = 187.57$, $p<0.0001$).

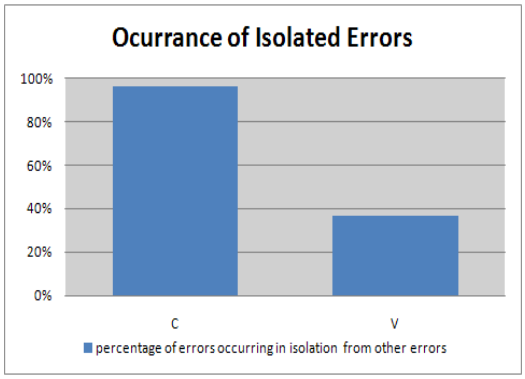
Errors on more than one contiguous segment did not coincide at the levels that would be expected if the errors were independent of each other and co-occurring at chance rates. Errors on both contiguous VC *and* CV sequences occurred more often than expected based on the rate of nucleus and coda errors.

Figure 4: Actual coincident error rates vs. chance.



Contiguous errors were not evenly distributed between segment types, either. Errors on Cs frequently occurred as isolated errors, while errors on Vs occurred most often in conjunction with errors on one or both adjacent consonants. See Figure 5.

Figure 5: Isolated errors on Cs vs. Vs.



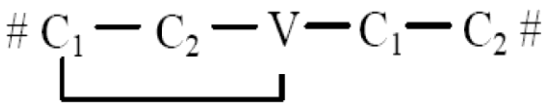
4. DISCUSSION

In answer to the main research question, there are in fact several asymmetries in the distribution of speakers' slips of the tongue across the syllable. Not only do speakers produce errors on vowels less frequently than on consonants, but errors in the nucleus position are less likely than in either the onset or coda position individually. The hierarchical model of the syllable (Fig. 1) does not predict this asymmetry, or provide an explanation for why it might occur, since the nucleus is as deeply embedded in the syllable structure as the

coda, and the model does not provide a structural basis for unifying all consonantal elements to predict robust and general differences between onset and coda consonants on the one hand, and the nuclear vowel on the other.

A model which could be used to predict the differential error rates is the Articulatory Phonology syllable model (Fig. 6)

Figure 6: The Articulatory Phonology syllable model.

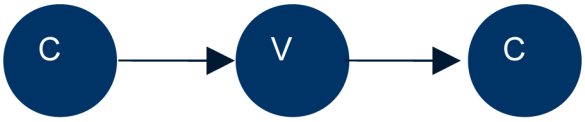


In the AP model, onset consonants are coordinated to be synchronous at their onset with the onset of the vowel gesture, and coda consonants are coordinated at their onset with the offset of the vowel. The vowel thus acts as an anchor to the rest of the syllable. This structure predicts that it will be more difficult for speakers to err when producing vowels without adversely affecting the production of the rest of the syllable.

Further support for this model comes from the finding that contiguous errors on both CV and VC pairs occur more frequently than by chance. The hierarchical model predicts a possible dependence between errors in the nucleus and coda (i.e., a dependence between errors that are in the syllable rhyme) or a possible dependence between errors that occur within the same syllable as a whole, but it does not predict any dependency relating errors in the onset and nucleus, since they do not form a constituent. In the AP model, however, the coordination between the vowel and both consonantal positions predicts why neither CV nor VC pairs behave independently with regard to errors. Finally, the position of the vowel as syllable anchor predicts the asymmetry observed between the rates of solo errors on Cs and Vs: most of the time that vowel errors do occur, there are other errors within the same syllable, which is not the case for consonants.

One finding the AP syllable model does not provide a clear prediction for is the higher frequency of errors in onset position compared to coda. A decrease in error rates as the syllable progresses is best predicted by a sequential-activation model of syllable construction (Fig. 7), in which each segment is activated by the previous segments, causing greater activation, and thus more accuracy (fewer errors) as the syllable progresses.

Figure 7: The sequential activation model.



To model all the findings from our study, a sequential activation element must be added to the Articulatory Phonology model of syllable structure. Sequential activation is consistent with the formulation of AP as a dynamical systems model in which the speech production system changes over the time course of production.

5. CONCLUSIONS

In a controlled study of speech production errors in American English, we confirmed the finding of previous studies of "caught" errors that speakers are less likely to make slips of the tongue when producing vowels than consonants. Additionally, we found several other asymmetries in speaker error rates, specifically that errors in onset position are the most likely and those in the nucleus the least; that consonant and vowel errors occur together more frequently than would be expected if they were independent, in both CV and VC pairs; and that although C errors are more likely to occur alone, V errors are more likely to occur in contact with C errors.

We propose that a model of syllable construction which can best predict all of these findings must include the properties of both Articulatory Phonology and sequential activation syllable models, which is consistent with a dynamical systems model of the speech production system.

6. ACKNOWLEDGEMENTS

Thanks to Drs. Gary Dell and Ryan Shosted, as well as all the members of the University of Illinois Phonotactic Learning Group for their helpful discussions and suggestions. This work is supported by NIH award HD044458 to Gary Dell, Cynthia Fisher and Jennifer Cole.

7. REFERENCES

[1] Browman, C.P, Goldstein, L. 1988. Some notes on syllable structure in articulatory phonology. *Phonetica*. 45, 140-155.

[2] Dell, G.S. 1984. Representation of serial order in speech: Evidence from the repeated phoneme effect in speech errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10(2), 222-233.

[3] Dell, G.S. 1988. The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language* 27(2), 124-142.
[4] Dell, G.S., Juliano, C., Govindjee, A. 1993. Structure and content in language production: A theory of frame constraints in phonological speech errors. *Cognitive Science* 17(2), 149-195.
[5] Goldstein, L., Pouplier, M., Chen, L., Saltzman, E., Byrd, D. 2007. Dynamic action units slip in speech production errors. *Cognition* 103(3), 386-412.
[6] McKay, D.G. 1970. Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia* 8(3), 323-350.
[7] Preacher, K.J. 2001. Calculation for the chi-square test: An interactive calculation tool for chi-square tests of goodness of fit and independence [Computer software]. *http://quantpsy.org*
[8] Shattuck-Hufnagel, S. 1979. Speech errors as evidence for a serial ordering mechanism in sentence production. *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*. Erlbaum.
[9] Stemberger, J.P. 1982. Syntactic errors in speech. *Journal of Psycholinguistic Research* 11(4), 313-345.
[10] Vousden, J.I., Brown, G.D.A., Harley, T.A. 2000. Serial control of phonology in speech production: A hierarchical model. *Cognitive Psychology* 41, 101-175.