Acoustic Correlates to Ambisyllabic Representations in American English

Monica Nesbitt

1 Introduction

Evidence for syllabic units in speech production is wanting. Much of the literature probing syllabic representations has been conducted in the realm of articulatory phonetics (Gick 2003, Krakow 1999, Turk 1994, etc.), only a handful of studies exploring these questions utilize acoustic methods. The few acoustic studies that have been conducted find only some evidence for temporal correlations to syllable affiliation (Byrd and Saltzman 2003, Coetzee and Wissing 2007, Selkirk 1982, 2001, etc.). Such correlations are often weak because, as Coetzee (2011) points out, articulatory differences between prosodic units are often very small. As such, the question remains: Are syllabic units part of a speaker/listener's mental reality? What information in the acoustic signal can be utilized to indicate syllabic affiliation? Can this information be used to shed light as to the syllabic structure of so-called ambisyllabic consonants?

Through an acoustic analysis of the Buckeye Corpus (Pitt et al 2007), this paper confirms that the syllable is indeed a part of a speaker's mental reality in American English. Consonants affect preceding vowels differently depending on their putative syllabic affiliation. Vowels preceding word-medial coda consonants are shorter and produced with a lower pitch than those preceding word-medial onset consonants. Likewise, speakers treat ambisyllabic consonants as though they are word-medial codas, producing vowels that precede them with a lower pitch and shorter duration than those that precede word-medial onsets. Pitch differences between the three consonant types were minor, while duration differences were significant. The results of this production analysis support recent perceptual findings which suggest that ambisyllabic consonants are treated as word-medial coda consonants by American English listeners.

1.1 Ambisyllabic Consonants

Ambisyllabic consonants have proven to be an anomaly in the theoretical as well as experimental literature. In this article, I will use the term *ambisyllabic* purely descriptively to refer to consonants, in American English, between vowels with primary stressed and an unstressed vowel, e.g. the /m/ in *lemon*. Three major analyses exist in the theoretical literature as to their syllabic structure: (a) such consonants are in onset position of the following syllable (Bermúdez-Otero 2013, Kiparsky 1979), (b) such consonants are in coda position of the preceding syllable (Borowsky 1986, Selkirk 1982, Wells 1990), (c) such consonants are multiply-linked to two syllables, occurring simultaneously as coda to the preceding and as onset to the following; i.e., they are ambisyllabic (Hayes 2009, Kahn 1976).

Proponents of the onset analysis appeal to the fact that intervocalic consonants are subject to the Maximal Onset Principle. This principle holds that as many consonants as possible be associated with the onset (Kenstowicz, 1994, pp. 257-258). Under this analysis, the /m/ in *lemon* must be in the onset of the second syllable. Proponents of the coda analysis, however, argue that while the Maximal Onset Principal is satisfied earlier in the derivation, stress assignment causes changes to syllable assignment. After stress is assigned, consonants in this particular environment, between a stressed and unstressed vowel, are resyllabified as codas on the surface (Selkirk 1982). This is due to the preference in American English for stressed syllables to be heavy. Under this analysis, *lemon* is first syllabified as le.mon and then resyllabified as lem.on on the surface. Proponents of the multiply-linked analysis also argue that stress assignment affects syllabic affiliation. For them, the phonetic realization of ambisyllabic consonants is used as evidence that they are indeed not purely onsets or purely coda consonants. The alternation of alveolar stops with flaps in American English is often cited as argument for multiply-linked representations. Alveolar consonants that occur in a

clear onset position are aspirated in American English, e.g., the /t/ in *attack*, when they occur in a clear coda position, however, they are glottalized, e.g., the /t/ in *atlas*. Crucially, when alveolar consonants are situated between a stressed and an unstressed vowel, alveolar stops are realized as flaps, e.g., the /t/ in *better*. For the multiply-linked proponents, this suggests that ambisyllabic consonants are something other than a pure coda or pure onset on the surface; they are both (Kahn 1976). The coda and multiply-linked analyses are both appealing because they allow these consonants to simultaneously satisfy the Maximal Onset Principle and the preference in American English for stressed syllables to be heavy, a consequence that is lacking in the onset analysis.

Just as there are disagreements in the theoretical literature, the experimental literature has not reached a consensus as to the syllabic representation of ambisyllabic consonants. Much of the experimental literature probing syllabic representations has used syllable-break tasks (cf. Gussenhoven 1986, Treiman and Danis 1988, Elzinga and Eddington 2014). While many of these studies have found some evidence for multiply-linked representations, some have found that syllable judgements are inconsistent across speakers (Elzinga and Eddington 2014). This is likely because participants are unsure of the task. The nature of these tasks relies heavily on a participant's assumed metalinguistic knowledge of syllable structure. As Steriade (1999) points out, it is possible that listeners in such experiments make word- or morpheme-edge judgements rather than syllable-edge judgments when they are explicitly asked to break words apart. For example, the training phase of the syllable-reversal task utilized by Treiman and Danis (1988) may have biased participants. For this task, participants were instructed to break up a word and repeat its parts in reverse order, e.g. *grandfather* would become *fathergrand*. The training phase for this task contained words like *grandfather*, where there was a morpheme or word boundary at the syllable break. So, it is not clear if participants were indeed breaking up monomorphemic words during the testing phase into syllables or if they were doing their best to break them up into possible English words or morphemes. As such, syllabic representation of ambisyllabic consonants is again unclear.

Using a less metalinguistic syllable-tracking task, Nesbitt and Durvasula (2016) found evidence for a coda representation of ambisyllabic consonants. For this experiment, the researchers aurally exposed participants to a list of words and instructed them to press a button on their keyboard when they heard a given sequence of sounds, e.g. /ham/ in that list of words. Crucially, one of the words in the list would be either a word whose first syllable shared the syllabic structure of the sequence, e.g., *hamlet,* or a word whose first syllable may match the sequences' syllabic structure, e.g., *hammock*. The expectation was that if listeners have a coda representation for ambisyllabic consonants, e.g. the /m/ in *hammock,* button presses when the word list contained *hammock* and when the list contained *hamlet* would be at the same rate when asked to track /ham/. If, however, ambisyllabic consonants have an onset representation, the rate of button presses for words with word-medial onsets, e.g. *locate,* and those with ambisyllabic consonants, e.g. *local,* would be the same. If, these consonants are indeed multiply linked, monitoring for words with ambisyllabic consonants should be different than that for onsets and codas. Ultimately, participants monitored syllables with a coda consonant, e.g., /ham/, at the same rate when that sequence occurred at the beginning of words containing word-medial coda consonants e.g., *hamlet*, as when it occurred at the beginning of words whose medial consonants were ambisyllabic, e.g., *hammock*. These same participants monitored sequences, e.g., /look/, at a different rate when the target words contained ambisyllabic consonants, e.g., *local,* than when they contained word-medial onset consonants, e.g., *locate*. From this, the authors concluded that ambisyllabic consonants have a coda and not an onset or multiply linked representation in American English. So, while some perception tasks support a coda analysis, others support a multiply-linked analysis and others find variable results.

This is also the case for previous production studies probing syllabic affiliation of ambisyllabic consonants. Some support a coda analysis (Krakow 1999, Turk 1994), and others have claimed that they are intermediate between onsets and codas (Gick 2003). As Durvasula and Huang (2017) point out, production studies often conflate *ambisyllabic* with *intervocalic*. Words containing intervocalic consonants that precede secondary stress; i.e., word-medial onsets, are often grouped with those containing intervocalic consonants that precede an unstressed vowel; i.e., ambisyllabic consonants. Separating these two consonant types is crucial, however, as onset consonants are argued to be different from ambisyllabic consonants.

During their investigation of nasalization effects on vowels preceding word-medial nasal consonants in American English, Durvasula and Huang (2017) controlled for the confounds

previously discussed. For them, word-medial onset consonants and ambisyllabic consonants were grouped separately and all target items were monomorphemic. They noted that at slow, normal, and fast speaking rates, ambisyllabic consonants patterned with word-medial coda consonants and not with word-medial onset consonant. There was no intermediate patterning of nasalization effects for ambisyllabic consonants. As such, the researchers argued that ambisyllabic consonants have a coda, not an onset or a multiply linked representation. If nasalization effects indicate syllabic affiliation, what, in the absence of a nasal consonant, do American English speakers utilize to indicate syllabic structure?

In this article, I will use two acoustic indices to probe the syllable structure of ambisyllabic consonants: (a) Vowel Duration, (b) Pitch. Though there are some inconsistencies with the production studies previously mentioned, many agree that vowels at an edge of a prosodic boundary are longer than when they occur within a prosodic boundary (Fougeron and Keating 1997). If the syllable is a prosodic unit, we would expect vowels in different syllabic environments to behave differently, such that those that occur at the edge of a syllable; i.e., preceding a word-medial onset, will be longer than those that occur within a syllable; i.e., preceding a word-medial coda consonant. Regarding pitch, a syllable with a coda consonant is produced with a lower fundamental frequency (pitch) than that of an open (CV) syllable (Lehiste 1970). I will compare the durational and pitch effects on the vowels that precede ambisyllabic consonants to those effects on vowels that precede word-medial onset and word-medial coda consonants to determine their syllabic affiliation in American English.

1.2 Crucial Comparisons

To probe syllabic affiliation, the following are the crucial comparisons that need to be made:

(1) word-medial onset consonants vs word-medial coda consonants
(2) word-medial onset consonants vs ambisyllabic consonants
(3) word-medial coda consonants vs ambisyllabic consonants

The first comparison will allow one to probe the acoustic effects of prosodic units in general, while the last two comparisons will help identify the syllabic structure of ambisyllabic consonants. If syllable structure is available in the acoustic signal, we should find that vowel measurements in words containing medial onsets are different from those containing medial coda consonants such that vowels before onsets are longer and produced with a higher pitch than those preceding coda consonants. If the current analysis finds no such differences for both vowel duration and pitch, then neither is a good index to probe the acoustic correlates of syllable structure, and suggests perhaps that syllable structure is not indicated in the acoustic signal, at least in naturalistic speech. Furthermore, if ambisyllabic consonants have a coda representation, their preceding vowel duration and pitch measurements should be similar to those preceding word-medial coda consonants. On the other hand, if they have an onset representation, their measurements should pattern with those of word-medial onset consonants. Finally, if their measurements are intermediate between onset and coda consonants, perhaps speakers in American English do indeed have a multiply-linked representation for these consonants.

2 Research Methods

For the current analysis, duration and pitch measurements of vowels preceding word-medial onset, word-medial coda, and ambisyllabic consonants were extracted from the Buckeye Corpus (Pitt et al 2007). Target words were controlled for primary stress position, word length and consonant type, such that only disyllabic words with primary stress on the first syllable whose word-medial consonant was a nasal, liquid or fricative were included in the analysis.[1] Likewise, all target items were monomorphemic, to avoid a confound with morpheme or word-boundary prosodic effects, as suggested by Durvasula and Huang (2017). A Praat script was utilized to automate the extraction

[1] Target words with word-medial glides, liquids and stops were excluded from the current analysis due to difficulties segmenting these segments in a speech stream.

of vowel duration and pitch measurements from the corpus, while plots and statistical analyses were conducted in R using the following packages: grepel, tidyr, gridEXTRA, stringi, dplyr and ggplot2.

2.1 Buckeye Corpus

Target items were extracted from the Buckeye Corpus (Pitt et al 2007). The corpus is a collection of audio recorded sociolinguistic interviews conducted in the late 1990's for The Ohio State University. Interviewees are Columbus, Ohio natives and are thus assumed to share the same dialect. The sample is stratified by age (under thirty and over forty) and by sex (male and female) and most speakers reported having upper-working to middle-class backgrounds. These interviews were casual in nature and topics ranged from politics to schools, traffic, and sports. The current analysis included interviews with 26 speakers (~30 minutes of speech per speaker).

2.2 Measurements

The Buckeye Corpus comes equipped with the .wav audio files of the interviews along with time aligned transcription files in .txt format. These files are annotated such that along with notes by the researchers, every word and phonological segment is parsed using the ARPAbet phonetic system. The .wav files and .txt grids were time aligned by the corpus personnel, such that the beginning and ending time of each segment and word was recorded on the .txt file. A script written in Praat was used to identify all of the vowels in each sound file. The script extracted the duration and pitch measurements (at the middle of the vowel) of every vowel in the Buckeye Corpus and kept track of the word in which they occurred.

Then, extracted duration and pitch values were normalized by speaker to account for variance in values between the speakers. The resulting measurements were narrowed down to only those vowels that preceded word-medial consonants in disyllabic words that had primary stress on the first syllable, as discussed in the next sub-section.

2.3 Target Word Identification

The CMU dictionary (Weide 1998) is a pronouncing dictionary of North American English. Like the Buckeye Corpus, the CMU dictionary is open source. It contains over 134,000 North American English words. This dictionary is especially useful, because like the Buckeye Corpus, it utilizes ARPAbet and marks its vowels for lexical stress to indicate how they are pronounced. After extracting all of the words in the dictionary, a script was used to narrow the list so that only disyllabic words with primary stress on the first syllable were included, for a total of 50,162 individual words. This list of words was used to identify the relevant words from the Buckeye Corpus.

For each identified word and annotation, the first vowel was isolated, then, the vowel duration and pitch measurements were extracted using the Praat script mentioned above. Each vowel was then assigned a condition according to which type of medial consonant appeared after it. Vowels that preceded consonants that were clearly codas of the first syllable (e.g. */l/* in *falcon*) were labeled *coda*, those which preceded a consonant and a syllable with secondary stress were labeled *onset* (e.g. /f/ in *seafood*), and those which preceded a consonant and an unstressed syllable were labeled *ambisyllabic*. In total, there were 10,037 vowels measured.

3 Results

The final list of target items included 10,037 individual tokens extracted from the Buckeye Corpus, for a total of 878 distinct disyllabic words with primary stress on the first syllable. Table 1 identifies the number of individual words and tokens represented in each of the three consonant conditions. As to be expected in American English, disyllabic words containing vowels with secondary stress are rare in the corpus. There are 8 times as many disyllabic words with ambisyllabic consonants and 5 times as many disyllabic words with coda consonants than there are words with onset consonants in the sample.

Target Type	Words	Tokens
ambisyllabic	423	5607
coda	345	3730
onset	110	700
Total	**878**	**10,037**

Table 1: Number of target items extracted from Buckeye Corpus.

As such, the current analysis will compare averages of the normalized vowel and pitch measurements in each condition. Table 2 displays the raw averages for these measurements.

Target Type	Duration(ms)	Pitch(Hz)
ambisyllabic	87.404	149.143
coda	92.470	147.677
onset	127.023	150.734

Table 2: Average duration (ms) and pitch (Hz) values for vowels in each condition.

Figure 1 displays the average normalized vowel duration measurements in each condition. Words with word-medial consonants are on the left in black, ambisyllabic words are in the middle in dark grey and words with word-medial onset consonants are on the right in light grey. Vowels preceding word-medial onset consonants are significantly longer than those preceding word-medial coda consonants (t = -2.5085, df = 20.509, p-value = 0.02062). Furthermore, vowels preceding ambisyllabic consonants are similar in length to those preceding word-medial coda consonants, (t = -1.9505, df = 61.975, p-value = 0.05564), but significantly shorter than those preceding word-medial onset consonants (t = -3.206, df = 20.409, p-value = 0.004354).

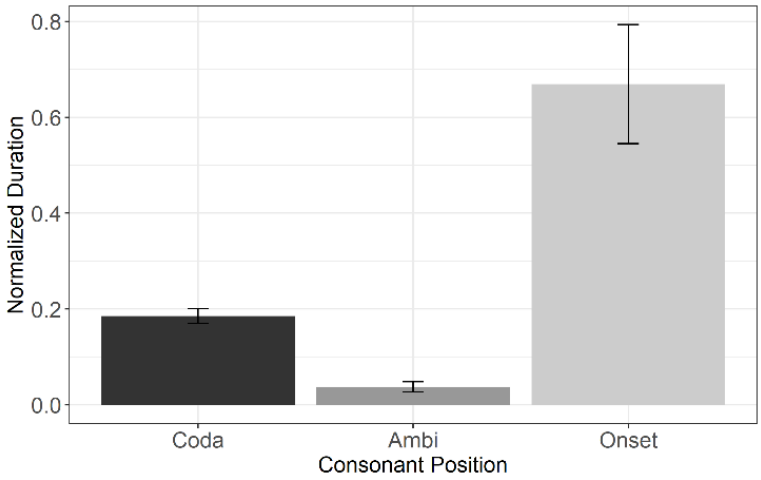

Figure 1: Average normalized duration measurements of vowels preceding three consonant types (word-medial coda, ambisyllabic and word-medial onset).

In Figure 2, we find a similar trend for pitch. American English speakers produce vowels preceding word-medial onset consonants at a higher pitch than those preceding word-medial coda consonants and ambisyllabic consonants, though these differences are not statistically significant: (t = -0.47668, df = 30.663, p-value = 0.637) and (t = -0.39997, df = 31.286, p-value = 0.6919), respectively. These speakers produce vowels preceding ambisyllabic consonants at the same pitch

level as they do vowels preceding word-medial coda consonants, (t = 0.10126, df = 61.959, p-value = 0.9197).

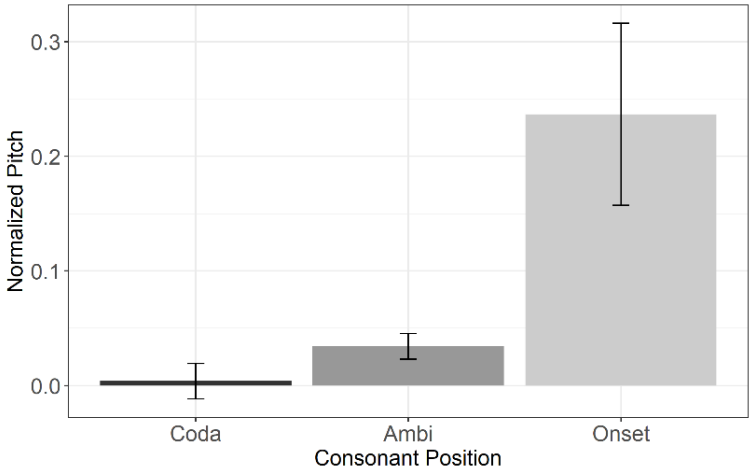

Figure 2: Average normalized pitch measurements preceding three consonant types (word-medial coda, ambisyllabic and word-medial onset).

4 Discussion

The current analysis set out to investigate the acoustic correlates to syllabic affiliation in hopes to determine the syllabic representation of ambisyllabic consonants. We were concerned with how word-medial ambisyllabic consonants patterned, as onsets, as codas, or as intermediate between the two. We focused on how these consonants effect duration and pitch of the vowels that proceed them, as these measurements have been fruitful in establishing differences between onset and coda consonants in the past. Using casual speech data from the Buckeye Corpus, the current analysis confirmed that word-medial onset consonants effect preceding vowels differently than word-medial coda consonants, such that vowels preceding onsets are longer and produced with a higher pitch than those preceding coda consonants. We also find that ambisyllabic consonants pattern with word-medial codas such that vowels preceding them are produced with a shorter duration and lower pitch than those preceding onset consonants.

As in previous studies (Durvasula and Huang 2017, Nesbitt and Durvasula 2016), the results of this study support a coda analysis of ambisyllabic consonants. The results are inconsistent with multiple-linkage representations, as these consonants did not affect preceding vowels in an intermediate fashion between word-medial coda and word-medial onset consonants. This may seem surprising given the alternation that occurs for alveolar stops in the ambisyllabic environment. As Huang and Durvasula (2017) point out, however, all of the different ambisyllabic analyses can account for flapping without having to rely on multiply linked representations. Onset and coda proponents posit that flapped realizations are simply the phonetic realization of these consonants and that these realizations have no bearing on their underlying representation (Kiparsky 1979, Selkirk 1982). The current study also simply describes the phonetic consequences of these consonants. I argue, though, that the phonetic realization can indeed tell us something about representations. If word-medial onset and word-medial coda consonants are realized differently, then phonetic realizations must tell us something about underlying syllabic representations.

Crucially, only a coda analysis can account for the fact that ambisyllabic consonants affect preceding vowels the same way that word-medial coda consonants do. A multiply linked analysis would have to explain why ambisyllabic consonants do not affect preceding vowels in an intermediate fashion between word-medial onset and coda consonants, a finding that does not coincide with the American flapping facts. An onset analysis would have to account for the fact that ambisyllabic consonants do not pattern with other word-medial onset consonants. The only

likely argument would rely, again, on stress assignment. Suggesting that stress has caused a change in syllabic affiliation. This is, however, an argument in support of the coda analysis, an argument that is born out in this paper.

Some may question the method of measuring preceding vowels in probing syllabic representations of the consonant. I note that when comparing the acoustics of the consonants in question, the results are the same. Durvasula and Felster (2017) found that at least for fricative and nasal consonants, ambisyllabic consonants are equal in length to word-medial coda consonants and not word-medial onset consonants.

The results of this experiment provide a few more avenues for further research. It is worth mentioning that while word-medial codas, ambisyllabic consonants and word-medial onset consonants exhibited differences in pitch effects, these differences were not statistically significant. In line with findings in previous studies, this suggests that while syllabic affiliation does affect pitch, the differences are very minor. The lack of difference between the consonant types is mostly due to the onset condition, however. The variance between speakers for word-medial onsets is what seems to be driving this result. It may be likely that speakers simply do not know how to treat secondary stress. This seems plausible as we see there is also considerable variation in the onset condition for the duration results. Perhaps examining multiple productions by individual speakers will help shed light on this hypothesis. Nonetheless, the results of this study suggest that measuring pitch alone is not fruitful in other studies probing prosodic boundaries. Preceding vowel duration, however, is a sound measurement for such comparisons. Likewise, Durvasula and Huang (2017) showed that nasalization is a sound measurement for probing syllabic affiliation, and Durvasula and Felster (2017) show that comparing duration measurements of the relevant consonants is also fruitful. Future studies are needed to elaborate on the acoustic properties of consonants at prosodic boundaries.

References

Bermúdez-Otero, Ricardo. 2013. An amphichronic approach to English syllabification. In *RUMMIT*.

Borowsky, Toni J. 1986. Topics in the Lexical Phonology of English. Doctoral dissertation, University of Massachusetts at Amherst.

Byrd, Dani, and Elliot Saltzman. 2003. The elastic phrase: modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics* 31:149-180.

Coetzee, Andries W. 2011. Syllables in speech processing: Evidence from perceptual epenthesis. In *Handbook of the Syllable*, ed. C. Cairns and E. Raimy, 295-328. Leiden: E.J. Brill.

Coetzee, Andries, and Daan Wissing. 2007. Global and local durational properties in three varieties of South African English. *The Linguistic Review* 24:263-289.

Durvasula, Karthik and Robert Felster. 2017. Syllabic affiliation of ambisyllabic consonants in American English. Paper presented at LSA 91, Austin, TX.

Durvasula, Karthik and Ho-Hsin Huang. 2017. Word-internal "ambisyllabic" consonants are not multiply-linked in American English. *Language Sciences* 62:17-36.

Elzinga, Dirk, and David Eddington. 2014. An experimental approach to ambisyllabicity in English. *Topics in Linguistics* 14:34–47.

Fougeron, Cecile, and Patricia Keating. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 10:3728–3740.

Gick, Bryan. 2003. Articulator correlates of ambisyllabicity in English glides and liquids. *Phonetic interpretation: papers in laboratory phonology* 5:222-236.

Gussenhoven, Carlos. 1986. English plosive allophones and ambisyllabicity. *Gramma* 10:119-141.

Hayes, Bruce. 2009. Syllabification in English. Ms., UCLA. URL http://linguistics.ucla.edu/people/hayes/251English/Readings/HayesAmbisyllabicity.pdf

Kahn, Daniel. 1976. Syllable-Based Generalizations in English Phonology. Doctoral dissertation, Massachusettes Institute of Technology.

Kenstowicz, Michael. 1994. *Phonology in Generative grammar*. Cambridge, MA: Blackwell.

Kiparsky, Paul. 1979. Metrical structure assignment is cyclic. *Linguistic Inquiry* 10:421–441.

Krakow, Rena. 1999. Physiological organization of syllables: a review. *Journal of Phonetics* 271: 23–54.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge: MIT Press.

Nesbitt, Monica and Karthik Durvasula. 2016. Ambisyllabic consonants are codas: Evidence from a syllable tracking task. Paper presented at LSA 90, Washington, D.C.

Pitt, Mark, Dilley, Laura, Johnson, Kenneth, Kiesling, Scott, Raymond, William, Hume, Elizabeth, and E. Fosler-Lussier. 2007. Buckeye Corpus of Conversational Speech 2nd release [www.buckeyecorpus.osu.edu] Columbus, OH: Department of Psychology, Ohio State University Distributor.

Selkirk, Elisabeth O. 1982. The syllable. In *The structure of phonological representations*, ed H. van der Hulst, and N. Smith, 337–383. Dordrecht: Foris.

Selkirk, Elisabeth O. 2001. On the phonologically driven non-realization of function words. Paper presented at BLS 27, University of California, Berkeley.

Steriade, Danca. 1999. Alternatives to syllable-based accounts of phonotactics. In *Proceedings of the 1998 Linguistics & Phonetics Conference,* ed. O. Fujimura, B. Joseph and B. Palek. Prague: Karolinum Press.

Treiman, Rebbeca and C. Danis. 1988. Syllabification of intervocalic consonants. *Journal of Memory and Language* 27:87-104.

Turk, A. 1994. Articulator phonetic cues to syllable affiliation: gestural characteristics of bilabial stops. In *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form*, ed P. Keating, 107-135. Cambridge, Mass: MIT Press.

Weide, Robert L., 1998. The CMU pronouncing dictionary. URL http://www.speech.cs.cmu.edu/cgibin/cmudict.

Wells, J. 1990. Syllabification and allophony. In *Studies in the Pronunciation of English, a Commemorative Volume in Honour of AC Gimson*, 76–86. London: Routledge.

Department of Linguistics and Languages
Michigan State University
East Lansing, MI 48824
nesbit17@msu.edu