Phonetic effects of onset complexity on the English syllable

Anna Mai
Department of Linguistics, University of California San Diego, La Jolla, CA, USA
acmai@ucsd.edu

Abstract: Although onsets do not arbitrate stress placement in English categorically, results from Kelly (2004) and Ryan (2014) suggest that English stress assignment is nevertheless sensitive to onset complexity and voicing. Phonetic work on languages in which onsets participate in categorical weight criteria shows that onsets contribute to stress assignment through their phonetic impact on the nucleus, primarily through their effect on nucleus energy (Gordon, 2005). Because they probabilistically participate in weight-based processes, it is predicted that onsets in English impact the phonetic realization of the syllable similar to the way that onsets do in languages with categorical onset weight criteria. To test this prediction, speakers in this study produced monosyllabic English words varying in onset complexity, and measures of duration, intensity, and pitch were collected. Results of the current study are consistent with the predictions of Gordon's perceptual account of phonological weight and further indicate that onsets impact properties of the intensity envelope itself, consistent with explanations for onset weight which appeal to onset influence on the perceptual center (Ryan, 2014). Given the similarity of these results to those found in languages with categorical onset weight, the findings reported here suggest that intrinsic phonetic effects of onset complexity are exploited differentially by languages with weight sensitive stress assignment.

Keywords: onset complexity, production study, p-center, syllable weight

1. Introduction

Crosslinguistically, onset-sensitive categorical weight criteria are rare (Davis, 1988; Downing, 1998; Gordon, 2005; Hajek & Goedemans, 2003; Hyde, 2007, 2; Topintzi, 2010; Topintzi & Nevins, 2017). However, the number and voicing of onset segments still predict syllable behavior in languages where onsets do not participate in categorical weighting criteria (Kelly, 2004; Ryan, 2014, 2018). For this reason, the probabilistic influence of onsets on weight-based behavior is argued to be phonetically motivated. This study investigates the phonetic motivations behind onsets' contribution to syllable weight in English through examination of their phonetic realization independent of their participation in weight-based processes such as stress assignment. By controlling for the presence of stress, this production study reports the intrinsic acoustic impact of onsets on syllable realization and argues that these effects are co-opted by the phonological weight system to enhance syllable prominence in American English.

In languages with weight-sensitive stress surveyed by Gordon (2002, 2007), syllable properties such as onset voicing, vowel length, vowel quality, coda complexity, and coda quality impact the placement of lexical stress. In these languages, syllables fall into weight categories based on some subset of these properties. Syllables considered heavier attract stress, while those considered lighter repel stress. Languages like Pirahã (KVV > GVV > VV > KV > GV) (Everett & Everett, 1984), which either draw more than three weight distinctions or draw weight distinctions based on onset properties are typologically rare. Gordon (2005) offers a perceptually motivated explanation for the observed typology: Given a phonemic inventory, weight categories will be drawn to maximize perceptual distinctiveness, where the phonetic property which has the greatest impact on perceptual distinctiveness is the energy (integrated intensity) of the nucleus. Since onsets have relatively small impact on nucleus energy, they have little impact on perceptual distinctiveness and consequently, rarely define weight categories.

From corpus and perceptual work on non-categorical onset sensitivity, Kelly (2004) and Ryan (2014) offer a complementary perceptual explanation for the impact of onset complexity on syllable weight. In languages with non-categorical onset sensitivity and weight-based stress assignment, the number and voicelessness of onset segments increases the likelihood of stress. In a corpus study of disyllabic English words, Kelly (2004) finds that the number of consonants in the initial onset of the word correlates positively with the probability of receiving initial stress. Ryan (2014) determines that this pattern is productive in English, finding that the number of consonants in the initial onset of disyllabic nonce words correlates positively with the probability that participants perceive initial primary stress. Ryan (2014) additionally shows that increased onset complexity draws the p-center (perceptual center; [Morton et al. 1976]) earlier in the syllable, allowing a greater portion of the syllable to be perceived as part of the rime, and suggests that onsets may contribute to syllable weight through their impact on the timing of the p-center.

We know from Gordon's work that weight can have acoustic origins, namely in integrated intensity. However, this has only been tested for categorical weight. Additionally, from Ryan's work, we know that onset complexity gradiently affects stress placement in English multisyllabic forms, and therefore, onsets should contribute to weight. Do complex onsets in English shift the phonetic realization of syllables in a way which could account for the perceptual results in Ryan (2014)? Previous work rightfully tested multisyllabic forms, since the authors were interested in stress placement. However, if the phonetic attributes of onset complexity were to be tested on multisyllables, then any effect found could be due to onsets, stress, or both. Using solely monosyllabic words to control for effects of stress placement would maximize the probability of observing results due to onset complexity alone. For this reason, this study uses English monosyllables with primary stress to examine solely the phonetic impact of onset complexity on the syllable, fleshing out phonetic properties of onsets independent of stress assignment which may be responsible for their probabilistic impact on weight-based processes in English.

In §2, categorical and gradient syllable weight are introduced and phonological factors which determine weight are discussed. Section 2.1 describes a perceptual account for categorical rime-based syllable weight and its extension to systems with categorical onset weight. Section 2.2 discusses gradient onset weight in English and describes an alternative, compatible perceptual account of weight criteria intended to account for gradient onset effects in English. Sections 3–5 describe and report the results of the production study, and §6 concludes.

2. Background

2.1. Categorical syllable weight

Syllable weight describes a language-specific division of syllables into two or more categories based on their segmental properties and their distribution with respect to a prosodic process such as stress assignment or tone licensing. These categories are referred to as "weights" and syllables are categorized as being of light, heavy, or intermediate weight (W. Allen, 1973; Hayes, 1989; Hyman, 1977, 1992, 2003; Jakobson, 1931; McCarthy & Prince, 1994; Trubetzkoy, 1939; Zec, 1988). Conventionally, syllable types which license the application of a process are called heavy while those which restrict the application of a process are deemed light. For example, in languages with weight-sensitive stress assignment, heavy syllables may be said to attract stress while light syllables may be said to repel stress. Crosslinguistically, intermediate categories may be assigned to account for languages whose prosodic processes exhibit more than two patterns of syllable behavior, but these systems are not as common as two-category systems (see descriptions in Gordon (2002), Hayes (1995), Ryan (2011)).

Across languages and processes, heavy syllables tend to be those which contain long vowels and more complex or sonorant codas, while light syllables tend to have short or centralized vowels and simple, obstruent, or absent codas (see survey in Gordon (2007)). However, as extensively documented in Gordon (2002, 2007), the division of syllables into weights is language-specific and is grounded in language-specific processes. Such processes exhibiting sensitivity to weight include word minimality requirements (McCarthy & Prince, 1994, 1995), reduplication (ibid.), compensatory lengthening (Hayes, 1989), tone licensing (Hyman, 2003), and stress assignment (Chomsky & Halle, 1968). Languages exhibiting weight sensitivity in the arbitration of one process may or may not exhibit sensitivity in the arbitration of another process, and syllable properties definitive of weight—also known as weight criteria—may differ across processes within the same language. For example, in Classic Greek, CVV syllables are considered heavy in poetic metrics, the assessment of minimal root requirements, and the assignment of pitch accents; CVC syllables are considered heavy in poetic metrics and the assessment of minimal word requirements; and CV syllables are considered light for all phenomena (Steriade, 1991).

The language– and process–specific nature of weight classes demonstrates that weight systems are not implemented deterministically from phonetic characteristics. A syllable's acoustic properties alone do not determine whether it will behave as light or heavy. In fact, Broselow, Chen, and Huffman (1997) argue the converse: phonological weight systems exert influence over the phonetic realization of different syllable types. Broselow et al. observe in Malayalam, a language in which CVV outweighs CVC and CV, that short vowels in closed syllables are significantly shorter than those in open syllables. In Hindi, a language in which CVV and CVC outweigh CV, they find no such difference in the length of vowels across open and closed syllables. For this reason, Broselow et al. argue that closed syllables in Hindi have a mora that those in Malayalam do not, using the phonetic distinction as evidence for difference in phonological structure.

Taken all together, these results suggest that while there exist crosslinguistic tendencies for strong syllables to exhibit greater rimal length and sonority than weak syllables, weight systems are not motivated purely by phonetics, and any adequate theory of syllable weight must incorporate both phonetic and phonological pressures into its account.

2.1.1. A perceptual account

Gordon (2002) presents a perceptual account of categorical weight criteria in which phonetic and phonological pressures are addressed by a simplicity metric. Citing Broselow et al.'s finding in Malayalam and Hindi, Gordon suggests that weight classes organize to maximize acoustic difference across simple weight criteria. Gordon's simplicity metric defines 'simple' and 'complex' criteria such that "a weight distinction is complex if it refers to >1 association between place predicates and weight units, or if it refers to disjoint representations of the syllable" (Gordon, 2002, p. 57). For example, a system in which only syllables with high, long vowels are heavy manipulates two features—height and length. Such a system is not attested crosslinguistically. Gordon further demonstrates that even when a complex structural distinction would outperform a simple distinction in terms of its phonetic distinguishability, languages do not make use of complex distinctions in their instantiations of weight criteria.

Alongside extensive documentation of other languages in which weight criteria align with language-specific points of maximal acoustic differentiation,[1] Gordon uses this simplicity metric to construct a theory of syllable weight in which the interplay of phonological processes and phonetic parameters together instantiate weight systems. By characterizing the optimal weight system as that which maximizes phonetic distinctiveness while minimizing the complexity of the phonological structures across which phonetic distinctiveness is assessed, Gordon outlines a theory of syllable weight which makes clear predictions about the weight criteria a language may instantiate given either phonological or phonetic data from which to predict. Given phonological data showing differential behavior of light and heavy syllables, Gordon's theory predicts maximal difference in perceptual energy (i.e. normalized integrated intensity) across the two weight categories. Given phonetic data showing integrated intensity across syllable types, Gordon's theory predicts that a weight distinction—if the language institutes one—would be drawn across the most phonetically distinct syllable types.

In formalizing a theory of syllable weight that draws upon phonetic effectiveness, Gordon examines two parameters of phonetic variation, duration and perceptual energy (i.e. loudness). He chooses these parameters for examination based on decades of evidence correlating duration, amplitude, and weight-based stress assignment across a variety of languages (Broselow et al. 1997; Duanmu, 1994; Ham, 2013; Hubbard, 1994, 1995; Maddieson, 1993; Zhang, 2001). Of the phonological processes exhibiting weight sensitivity, stress assignment is extensively documented and demonstrates the most diversity of weight criteria. For this reason, Gordon (2002) fleshes out his phonetically driven account of syllable weight using evidence from stress assignment.

2.1.2. Extension of the perceptual account to categorical onset criteria

Gordon (2005) addresses onset-sensitive weight criteria within the framework proposed in Gordon (2002). Initially, in this framework, onsets appear hard pressed to meet the criterion for phonetic effectiveness due to the relatively marginal contribution that onsets make to a syllable's perceptual energy. However, as Gordon (2005) shows, in languages exhibiting onset-sensitive weight criteria it is possible for onset-based distinctions to outrank rime-based distinctions with respect to phonetic effectiveness. For example, in Eastern Arrernte two syllable words bear initial stress irrespective of their syllable structure, while words of three or more syllables bear initial stress only if the initial syllable has an onset and irrespective of rimal properties (Davis, 1988; Goedemans, 1998; Strehlow, 1942; Topintzi & Nevins, 2017). Examining the phonetic effectiveness of Arrernte's weight system, Gordon finds the greatest difference in mean perceptual energy between onsetless and onsetful syllables, and upon further analysis attributes the phonetic efficacy of the onset distinction to the fact that the vocalic portion of onsetless syllables was shorter than that of syllables with onsets.

Having shown the way in which onset-sensitive weight impacts the phonetic realization of the syllable beyond the properties of the onset itself, Gordon (2005) appeals to auditory nerve adaptation as a possible physiological basis for his finding. Auditory nerve fibers (ANFs) fire maximally in response to rapid changes in frequency and amplitude that occur within their characteristic frequency (CF) bandwidth. An ANF with a low CF may respond preferentially to transitions from obstruent to sonorant segments or at the onset of voicing or nasality, responding to the rapid increase in lower frequency power. Similarly, an ANF with a high CF may respond preferentially to transitions from sonorant to obstruent segments or at the onset of stop release, responding to the rapid increase in higher frequency power (Delgutte, 1997). In this way, the ensemble activity of high and low CF cells is capable of encoding fluctuations in sonorancy necessary to compute the structure of syllables or syllable-like structures (cf. the interval; Steriade (1997, 1999)).

However, the aspect of ANF speech encoding relevant to Gordon's theory of onset-sensitive weight is that of adaptation. Adaptation of the auditory nerve plays several roles in speech coding (Delgutte, 1982, 1986; Delgutte & Kiang, 1984), key among them an increase in the temporal resolution of onset representation and an enhancement of differences in spectral content across contiguous time windows. Over the course of a period of invariance the firing rate of an ANF gradually decays, 'adapting' to the stimulus. In part this adaptation may be caused by depletion of neurotransmitter at the synapse between the cochlear hair cell and the ANF, but also may result from center-surround suppression of the ANF receptive field. The center-surround organization of ANF receptive fields operates such that ANFs of similar characteristic frequency (CF) to the

ANF activated by a stimulus inhibit the activity of the activated ANF over time. When a change occurs in the stimulus, the suppression of 'adapted' nerve fibers increases the signal to noise ratio of the 'fibers' whose CF are tuned to fire maximally to the new stimulus. In this way, ANF adaptation increases the temporal resolution of shorter, transient events (e.g. onsets) and enhances information-rich periods of transition in the speech signal.

As it pertains to a perceptual account of onset contribution to syllable weight, Gordon (2005) argues that weight best corresponds to percepts of loudness which in turn correspond to syllables during which ANFs fire at a higher rate. With this understanding of the neural substrate for syllable weight, the mechanism of auditory adaptation predicts three observations that bear out in onset-sensitive weight systems crosslinguistically: 1. more complex onsets prolong the adaptive decay of ANF activity resulting in greater enhancement of the ANF response at the transition into the following vowel; 2. lower sonority onsets whose amplitude and spectral content differ greatly from the vowel which follows them will elicit greater ANF activity and thus will carry greater weight; and 3. onset weight criteria are rare crosslinguistically because long vowels and coda material dwarf onsets in their overall energy such that any boost in ANF firing rate afforded by onset characteristics make an ineffective contrast for the purposes of weight criteria.

2.2. Gradient onset weight

Although categorical onset weight criteria are rare, onsets are capable of influencing weight systems, even in systems that do not reference onset characteristics in categorical weight criteria. A series of behavioral studies in English and corpus studies examining onset-sensitive weight across several additional languages show gradient effects of onset characteristics on weight assignment (Kelly, 2004; Ryan, 2014). Admittedly, as shown by Nanni (1977), English adjectives formed with the suffix *–ative* demonstrate categorical onset-sensitivity, where the first vowel of the suffix will receive secondary stress if its onset is filled by an obstruent and will receive no stress if filled by a single sonorant, i.e. innovative /ˈɪnəˌveɪtɪv/ vs. manipulative /məˈnɪpjələtɪv/. However, beyond this subset of the English lexicon, onset criteria are not evidenced in the weight systems investigated by Kelly (2004) and Ryan (2014), yet the gradient impact of onset characteristics on phonological processes in these languages is pervasive: onset characteristics influence the probability of stress assignment in English, Russian, and Italian, and they influence the probability of a syllable being placed in a metrically strong position in English, Sanskrit, and Finnish.

Based on a corpus of two syllable words, Kelly (2004) shows that the probability of stress assignment to the first syllable of a two-syllable word increases as the number of segments contained in the onset increases. Furthermore, when Kelly presented English speakers with written pseudowords and asked them to assign stress to the first or second syllable, the number of segments in the onset similarly increased the probability of stress assignment to the first syllable. Examining 62 pairs of monosyllabic rhyming words found in Milton's Paradise Lost, Kelly also found that words of greater onset complexity were more likely to be found in strong metrical positions. Monosyllabic rhyming pairs were matched such that they differed only in onset complexity, demonstrating that onsets not only influence weight relative to other syllables within a word, they also influence weight independent of within-word comparison, as monosyllables.

Ryan (2014) extends the breadth of Kelly's studies to non-English languages and meters and extends their depth within English to examine the impact of onset voicing. Using a subset of the CELEX corpus, Ryan corroborates results from Kelly (2004) and further shows that voiceless stop onsets attract weight more than voiced stops and that two-segment onset clusters which do not contain /r/ attract more weight than those that do. These two findings accord with typological observations made by Gordon (2005) as well as predictions made by Gordon's perceptual account of onset weight, compatible with his auditory adaptation account of syllable weight. Gordon (2005) notes the crosslinguistic attribution of greater weight to less sonorous onsets and makes the typological generalization that when a weight distinction is made by onset featural content, lower sonority segments will be treated as heavier. For this reason, Gordon posits that a phonetic correlate of obstruence may play a role in the determination of onset weight. For Gordon and the auditory adaptation account, this correlate may be as simple as the duration of stop closure (the period of greatest amplitude attenuation), where the intensity of the onset plays the primary role in influencing syllable weight, and less intense onsets are predicted to contribute greater weight.

However, Ryan (2014) observes one additional phonetic correlate of onset complexity in his study which leads him to attribute onset weight to the p-center account rather than the theory of auditory recovery proposed in Gordon (2005). When examining the relationship between syllable duration and onset complexity, Ryan looked at initial, stressed, open syllables of multisyllable words in a subset of the Buckeye Corpus and found that vowel duration decreased as onset complexity increased. This finding suggested that the rime alone could not be the phonetic locus of onset weight effects. If vowels tended to be longer following more complex onsets, the simplest explanation for onset weight would be its impact on the rime, and measurements primarily driven by the rime—such as Gordon's integrated intensity—would provide the most parsimonious explanation for the data. However, if increased onset complexity tends to decrease the duration of the vowel, onset weight likely should not be attributed solely to rimal duration or intensity. For this reason, Ryan attributes onset weight to an account in which onset properties influence the domain of weight evaluation, namely the p-center account.

2.2.1. A perceptual account for gradient weight criteria

Ryan (2014) proposes that the percept of syllable weight corresponds to the *perceptual center* (hereon p-center) of the syllable, a concept developed by Morton, Marcus, and Frankish (1976) from that of the 'syllable beat' (G. Allen, 1972; Rapp-Holmgren, 1971) to address the fact that the perceptual isochrony of a series of words is incongruent with several different measures of acoustic isochrony, including isochrony of word onset, isochrony of stressed vowel onset, and isochrony of the position of peak intensity of the stressed vowel (Morton et al. 1976, p. 406). Morton et al. (1976) define the p-center of a (monosyllabic) word as its "psychological moment of occurrence," (p. 405) and within a series of words, it is the point during a word which must be regularly timed for isochrony to be perceived. In the most common type of task to determine a word's p-center, a participant hears a series of alternating sounds (i.e., 'base' words and 'test' clicks) and is asked to adjust the rhythm of the sounds (Cooper, Whalen, & Fowler, 1986; Harsin, 1997; Pompino-Marschall, 1989, among others). While the interval between the 'base' sounds is fixed, the participant can adjust the timing of the 'test' sounds with a controller, and they are asked to adjust the timing of 'test' sounds until they feel the 'test' sounds are synchronous with the 'base' sounds. The p-center is then defined as the relative time during a 'base' sound that the attack of the 'test' sound occurs.

The relationship between the p-center and syllable weight was first observed by Browman and Goldstein (1988), who noticed a qualitative similarity between the behavior of the p-center and the C-center, or *consonant center*, defined as the mean of the articulatory midpoints of a sequence of consonantal gestures. As onset duration and complexity increase, the C-center and p-center both occur later in the syllable (Browman & Goldstein, 1988; Rapp-Holmgren, 1971), and both maintain a consistent duration from the closure of a syllable-final stop, regardless of onset complexity (Fowler & Tassinary, 1981). For this reason, Browman and Goldstein (1988) speculate that the p-center, like the C-center, may be "a universal syllable-initial metric" for syllable weight.

Goedemans (1998) offers a perceptual account for syllable weight which centers perceived duration. Rather than grounding the acoustic correlate of syllable weight in the perception of intensity, Goedemans (1998) suggests that syllable duration influences the perception of syllable weight. Thus, if the p-center marks the beginning of a syllable's duration relevant for weight criteria, then the earlier p-center associated with more complex onsets should result in greater attribution of syllable weight all else being equal. For example, even if the rime of *rain* was produced identically to the rime of *train*, the additional onset consonant in *train* would advance the p-center within the word, such that the domain of weight evaluation in train would be greater than the domain of evaluation in rain. If weight were evaluated purely by the duration of this domain, train would be heavier than rain.

Using the syllable intensity maximum as a proxy for the p-center, Goedemans (1998) attempts to manipulate the perception of syllable duration through manipulation of the location of the intensity maximum. Using these methods Goedemans finds no difference in the perception of syllable duration and concludes that intensity maxima have no bearing on the determination of syllable weight. However, this conclusion assumes that 1) duration is the primary acoustic cue to syllable weight and 2) the p-center determines the portion of the syllable available for the perception of duration and by extension, for the perception of syllable weight. Given the results obtained in Gordon (2002) showing that acoustic correlates of loudness best predict weight, the poor explanatory power afforded by measurements of duration and perceived duration come as no surprise. Like Gordon, Goedemans shows that duration does not predict syllable weight well. However, Goedemans' finding alone does not preclude the possibility of the intensity maximum 1) being affected by onset characteristics or 2) playing a role in the determination of syllable weight. Furthermore, models of the p-center which incorporate percepts of loudness more successfully predict the location of the p-center (Villing, 2010), suggesting possible convergence of both p-center and perceptual energy accounts of syllable weight. For this reason, the current study treats acoustic correlates of the p-center (pitch and intensity maxima) as additional factors in the perceptual energy account advocated by Gordon.

3. Production study

Previous acoustic studies of onset weight have used multisyllabic words to demonstrate the role onsets play in stress assignment. Using multisyllabic words allows RMS amplitude to be normalized against another syllable within the word domain Gordon (2002, 2005, 2007) and allows the probability of stress assignment to be measured relative to another syllable position within the word domain (Kelly, 2004; Ryan, 2014). These studies cleanly show that onset complexity and onset voicelessness increase the likelihood that a syllable will receive stress, and both accord with the perceptual theory of syllable weight presented in Ryan (2014) to account for the gradient effect of featural and segmental properties of the onset on categorical syllable behavior.

This study builds off the results of Kelly (2004) and Ryan (2014) to determine the extent to which the phonetic correlates of onset weight in American English are correlates of onsetful syllables more generally, and the results of this study have implications for the perceptual accounts of onset weight advanced in both Gordon (2002, 2005) and Ryan (2014). Gordon (2005) and Ryan (2014) concur that phonetic properties of onsets contribute to their impact on categorical weight assignment. However, neither adjudicates whether onsets exert influence on stress assignment due to general acoustic properties of onsets in the language, or if the acoustic properties of onsets are enhanced solely to the advantage of stressed syllables, playing a role in stress assignment through acoustic difference relative to onsets of unstressed syllables.

To show the stress-independent impact of onsets on the acoustic realization of the syllable, monosyllabic words are used. Since monosyllabic words offer only one location for stress placement, stress assignment in monosyllabic content words is trivial. In this way, the use of monosyllabic content words controls for the impact of stress on the realization of the syllable.

Controlling for stress is crucial to distinguish the acoustic properties of stress from the acoustic properties of onsets themselves and has not been directly addressed in previous literature. Gordon's studies of languages with categorical onset weight criteria do not mediate between these two possibilities because they treat only multisyllabic words in languages with onset weight criteria. Ryan's work on English does not disentangle intrinsic onset properties from those conditioned by stress because phonetic properties are assessed against the probability of receiving stress, a metric which pools stressed and unstressed syllables. Kelly (2004) comes closest to differentiating intrinsic from stress-conditioned onset properties in the paper's third experiment, showing that even monosyllables demonstrate onset-sensitivity in weight phenomena like metered verse. Kelly's observation is crucial because it demonstrates that onset-sensitive stress assignment, while gradient and probabilistic in its application, is a property which is absolute within the syllable domain; onsets exert influence on weight independent of comparison against adjacent syllables.

For this reason, the data presented in the current study comprise single syllable words. The use of single syllable words controls for the possibility that onset-sensitive weight is calculated relative to unstressed syllables by circumventing the assumption that stressed and unstressed syllables share a common pattern of acoustic realization. In this way, collection of data which controls for stress may yet inform questions relevant to the study of stress and its weight-conditioned assignment. Using this experimental design, the current study shows that 1. the acoustic realization of onsets in an onset-sensitive stress language parallels acoustic results reported in languages with categorical onset weight criteria and 2. The acoustic measures varying across onset complexity correspond to those which were found to correlate with propensity to receive stress in multisyllabic words. These results suggest that the properties of onsets observed in this study reflect onset properties occurring generally in the English language.

In the current study, English speakers pronounced one-syllable words of varying onset complexity and measurements of intensity (peak & average), duration (normed onset, normed rime, & total), and f0 (peak & average) were taken. While previous work on English has described phonetic effects of onset complexity on stressed syllables in multisyllabic words (Ryan 2014), this study extends the description of onset properties to single syllable words, demonstrating onset properties independent of other syllables within the lexical domain.

3.1. Participants

Nineteen native English speakers (12 female) were recruited through the SONA UC San Diego Experiment Scheduling system. Participants were 17-22 years of age, all began speaking English before the age of six, and nine specified learning English in California. Two participants indicated learning Vietnamese before the age of six in addition to English, and three indicated learning Spanish before the age of six in addition to English.

3.2. Materials

Participants spoke aloud the phrase "Please say X again," where X was replaced with one of 160 target words or 40 filler words, for a total of 200 utterances recorded per participant. All items were spoken in this carrier phrase to control for effects of variable utterance position and coarticulation across word boundaries. Target items were real monosyllabic words of English, all of which contained one coda, one vowel, and between zero and three onset segments (i.e. *ate, sate,*

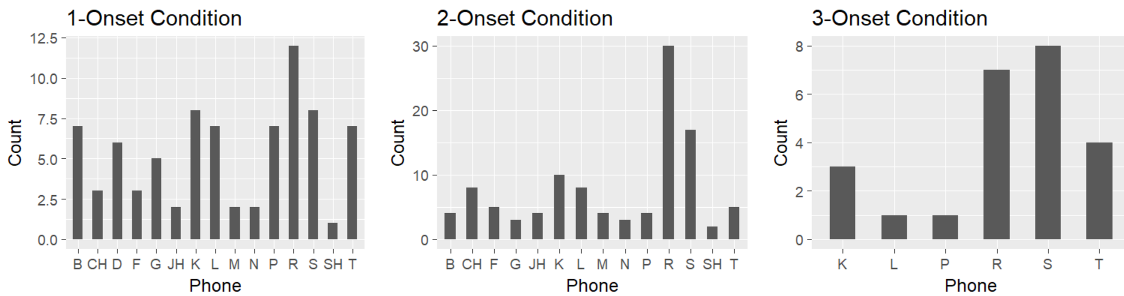

Figure 1 Target stimuli onset phone frequency across onset complexity conditions. Phones are labeled using ARPABET phonetic transcription codes.

state, straight). Target words fell into one of 8 vowel categories and one of 20 rime categories, as shown in Table 1. Each target word was accompanied by at least one other target word in the same rime category but differing in onset complexity (i.e. *rain, train*), as shown in the righthand column of Table 1. Using solely monosyllabic words as targets in this study controls for effects of stress placement, maximizing the probability of observing results due to onset complexity alone. Filler items were real two-syllable words of English carrying stress on the initial syllable (i.e. *splendor, market*). These fillers were designed to perturb the prosodic monotony of target items, reducing the chance that a participant would produce the carrier phrases with any kind of list intonation. A full list of target items and fillers is included in the Appendix.

Target items are not fully balanced across all onset, nucleus, and coda types, but care was taken to avoid bias across complexity conditions in segmental features which had the potential to skew results. Segments known to have longer intrinsic length were balanced as best as possible with those known to have shorter intrinsic length. For example, among the eight nucleus types represented in the target items, there are two front vowels (/æ/, /ɪ/ (1 high)), two back vowels (/ɑ/, /u/ (1 high)), three diphthongs (/oʊ/, /ɑɪ/, /eɪ/), and central schwa. The front vowels are contained in six rhyme types, the back vowels and schwa account for another six rhyme types, and diphthongs account for nine rhyme types. Additionally, twelve different coda types are represented across the twenty rime types. Nine out of the twenty rime types contain stop codas, and ten out of the twenty rime types have voiced codas. Eleven different rime types are represented among zero-onset words, and seven rime types are represented among three-onset words. Four of these rime types are shared across all four onset conditions.

Figure 1 shows the distribution of phonemes in onset position across the three onsetful conditions. Phonemes /ɹ/ and /s/ are overrepresented in the two-onset condition, but one– and two–onset conditions otherwise contain comparable distributions of sounds. Target items in the three-onset condition exhibit less phonemic diversity due to the phonotactic constraints of English.

3.3. Methods

Monophonic recordings were taken in a sound attenuated booth in the UCSD Phonetics Lab, collected using Praat software (Boersma & Weenink, 2011) and a preamplified head-mounted microphone. Materials were presented token-by-token to participants in random order using PsychoPy visualization packages (Peirce, 2007). Each screen depicted the carrier phrase “Please say X again,” written in standard English orthography, where the ‘X’ was replaced with a target word. Participants were instructed to read each phrase aloud at a comfortable pace.

Table 1 Target items. A representative subset of target items for each vowel and rime used in the study.

Vowel	Rhyme	Example
æ	/æp/	rap, trap
	/æʃ/	ash, rash
ɪ	/ɪl/	ill, shill
	/ɪŋk/	ink, sink
	/ɪm/	rim, trim
	/ɪp/	rip, trip
u	/ud/	rude, crude
ə	/əm/	rum, drum
	/əg/	rug, shrug
ɑ	/ɑt/	ought, taught
	/ɑf/	cough, scoff
oʊ	/oʊθ/	oath, both
	/oʊk/	oak, soak
aɪ	/aɪl/	isle, mile
eɪ	/eɪp/	ape, cape
	/eɪt/	ate, Kate
	/eɪl/	ale, kale
	/eɪm/	aim, fame
	/eɪn/	rain, train
	/eɪk/	ache, rake

4. Analysis and results

Target words were segmented and labeled using Praat TextGrids. Segmentation was determined using visual landmarks in the spectrogram. The onset of the second formant marked the beginning of the vowel, and the offset of the second formant marked the end of the vowel. Word-initial stops were segmented at the point of release and at the onset of the vowel. The beginning and end-point of word-final oral stops was considered the offset of the preceding vowel, which is to say, word-final oral stops were not segmented. Fricatives were considered to start at the onset of frication noise and were considered to end at the offset of frication noise or the onset of the following vowel if pre-vocalic. The onset of nasal stops was marked by a decrease in intensity and the onset of the lowest nasal formant. The offset of a nasal stop was marked by a visible increase in intensity if pre-vocalic or the offset of the lowest nasal formant if word-final. The onset of other sonorant consonants (/ɹ/ and /l/) was marked by a decrease in intensity, and the offset of these sounds was marked by a subsequent increase in intensity.

Following segmentation, segments were coded by their structural position in the syllable (i.e. onset, nucleus, or coda), and segment and total word durations were measured in Praat. Total onset duration and total rhyme duration were also calculated by summing the durations of segments in the onset and segments in the rime, respectively. Segment, onset, and rime durations were then normalized relative to the total word duration through division by the total word duration. This measure resulted in normalized durations between 0.0 and 1.0, representing the portion of the word occupied by each segment (or onset or rime). Any normalized measure outside this range was considered the result of measurement error, and such tokens were regarded as defective and discarded for subsequent analysis.

An f0 contour was extracted from each word using Praat's autocorrelation algorithm to detect acoustic periodicity (Boersma, 1993). To set the pitch floor and ceiling parameters adequately, speakers were impressionistically binned as having 'low' or 'high' voices. A pitch floor of 75Hz and pitch ceiling of 300Hz were used for 'low' voices, a pitch floor of 120Hz and ceiling of 500Hz were used for 'high' voices, and the algorithm sampled each signal at a rate of 100Hz. From this f0 contour, peak and average f0 were calculated for each word in Hertz, and average f0 was also calculated for each sonorant segment. For each word, the timing of the peak f0 was calculated relative to the beginning of the word by dividing the time point at which the peak f0 occurred by the total duration of the word. This measure resulted in normalized times between 0.0 and 1.0, representing the portion of the word preceding the time of the word's f0 maximum. Any normalized measure which fell outside this range was considered an error, and such tokens were discarded for subsequent analysis.

An intensity contour was extracted from each word using Praat's default algorithm which squares the power spectrum at each time step and convolves it with a Gaussian window. The sampling rate for this computation was set at 100Hz. From the resultant intensity contour, peak and average intensity (dB) were calculated for each word, and average intensity was also calculated for each segment. For each word, the timing of the maximum intensity within each word was calculated relative to the beginning of the word by dividing the time point at which the peak f0 occurred by the total duration of the word. This measure resulted in normalized times between 0.0 and 1.0, representing the portion of the word preceding the time of the word's intensity maximum. Any normalized measure which fell outside this range was considered an error, and such tokens were discarded for subsequent analysis.

Additionally, each sample of the intensity contour and its corresponding time point were logged for each word. These values were then treated as a time series from which the finite derivative was calculated for each time point and convolved with a Gaussian filter to smooth the result. This procedure resulted in a time series capturing the rate at which intensity changed over the course of the word. From this time series, the time point at which the intensity was maximally changing and the value of its slope was logged for analysis described in §8.3.

Throughout the analysis, the Kruskal-Wallis H test (Kruskal & Wallis, 1952) and Dunn's post hoc (Dunn, 1964) with Bonferroni correction for multiple comparisons are used in lieu of the standard one-way ANOVA and Tukey's post hoc due to the data's non-normal distribution. The Kruskal-Wallis H test is a non-parametric method of testing for stochastic dominance between groups. It assumes that samples are independent and come from similarly shaped distributions, but it does not require that samples be normally distributed nor of the same size (Kruskal & Wallis, 1952). Non-parametric analysis of variance methods, like the Kruskal-Wallis test, are conservative relative to their parametric counterparts, and in a review of nonparametric range tests including Dunn's post hoc test for significance, Day and Quinn (1989) find that the rate of Type 1 errors when using Dunn's post hoc is low, even when comparing groups with heterogenous variances or differing sample sizes. For this reason, one can be reasonably confident that the significant differences between samples reported in the following analyses are not erroneous.

4.1. Duration

A Kruskal-Wallis test was conducted to compare the effect of onset complexity on word duration in zero, one, two, and three onset segment conditions. There was a significant effect of the number of onset segments on word duration at the $p < .0001$ level for the four conditions ($\chi^2 = 209.58$, $p < .0001$). Post-hoc comparisons using Dunn's test indicated that each onset complexity condition differed significantly from all other complexity conditions at $p < .0001$. Figure 2 plots this result.

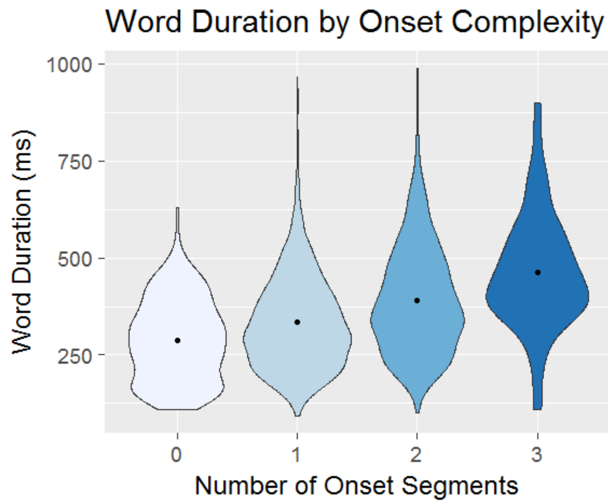

Figure 2 Raw word duration increases with the number of segments in the syllable onset. Black dots indicate mean values.

The positive relationship between raw word duration and onset complexity is best explained by the greater number of segments in words of greater onset complexity. Rather than reflecting a property specific to onsets, the result shown in Figure 2 likely reflects the fact that many segments take more time to produce than few segments. However, when we examine normalized measures of the onset and vowel duration, a more precise image of onset complexity's impact on duration emerges.

A Kruskal-Wallis test comparing the effect of onset complexity on the proportion of the syllable occupied by the onset shows a significant effect of the number of onset segments on the normalized duration of the onset ($\chi^2 = 781.98$, $p < .0001$). Post hoc comparisons using Dunn's test indicate that each onset complexity condition differs significantly from all other conditions at $p < .0001$. All together these results show that as onset complexity increases, so does the proportion of the syllable occupied by the onset (Fig. 3).[2] Conversely, a Kruskal-Wallis test comparing the effect of onset complexity on the proportion of the syllable occupied by the vowel shows a significant effect of onset complexity on the normalized duration of the vowel in the opposite direction ($\chi^2 = 442.00$, $p < .0001$). Dunn's post hoc comparisons indicate that each onset condition differs significantly from all other conditions at $p < .0001$. As onset complexity increases, the proportion of the syllable occupied by the vowel decreases. These effects are summarized in Figures 3 & 4.

The durational results presented here accord with those reported by Ryan (2014). When Ryan examined a set of word-initial, stressed, open syllables, he found that onset complexity and vowel duration exhibit an inverse, trading relationship. As onset complexity increases, vowel duration decreases. Ryan uses this result as evidence against the hypothesis that onsets contribute to syllable weight solely through their impact on the vowel. Results of the current study corroborate this argument, replicating results found in Ryan (2014) and in the wider literature on syllable compression (Browman & Goldstein, 1988; Katz, 2010). From the perspective that the rhyme is the arbiter of weight both perceptually as well as phonologically, a proportionally smaller rhyme ,in syllables with complex onsets is unexpected, and it is for this reason that Ryan (2014) sought explanatory power for a perceptual account of onset-sensitive weight from the p-center. Like Ryan's evidence from open syllables, the current data from closed syllables neatly suggest a role

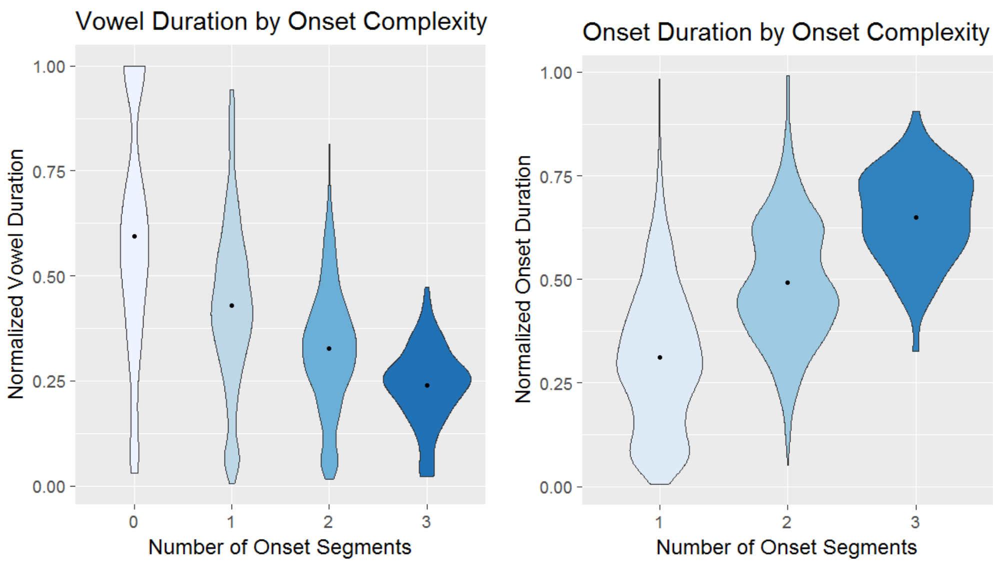

Figure 3 The proportion of the syllable occupied by the onset is greater in words with more onset segments (left), while the proportion of the syllable occupied by the vowel is lesser in words with more onset segments (right). The normalized duration values plotted on the y-axis of each figure correspond to the duration of the onset (left) or rime (right) divided by the total word duration, respectively. Black dots indicate value means.

for acoustic correlates other than rhyme duration in the assessment of onset contribution to syllable weight. Data from the current study rule out the possibility that absolute rhyme duration alone drives the influence of onsets on closed-syllable weight.

4.2. Timing of f0 and intensity maxima

The timing of the word's intensity maximum was calculated relative to the beginning of the vowel by subtracting the normalized time at which the intensity maximum occurred from the normalized duration of the onset. Similar to a measure like VOT, which indicates whether vocal fold vibration begins before or after the release of a plosive, this measure of the intensity maximum's timing indicates whether the intensity maximum occurs before or after the beginning of the vowel (the end of the onset). Rather than examining the normalized time of the intensity maximum relative to the beginning of the word (using the metric described in §6), the intensity maximum was calculated relative to the beginning of the vowel because the vowel is most likely to contain the amplitude maximum. Of all segments in the word, the vowel is the most sonorant, produced with the least constriction of the vocal tract. For this reason, it has the least attenuated intensity, and as such, it likely contains the intensity maximum. Since the duration of the onset increases with onset complexity, measuring the timing of the intensity maximum relative to the beginning of the word would overwhelmingly reflect the onset's duration rather than any shift in the intensity maximum's timing within the domain that it is most likely to occur (the vowel). For this reason, the timing of the intensity maximum was evaluated relative to the beginning of the vowel. Figure 5 provides a representative example of how this measure was calculated. From this point forward, any reference to intensity maximum timing will refer to this measure.

A Kruskal-Wallis test comparing the effect of onset complexity on the timing of the word's intensity maximum shows a significant effect of the number of onset segments on the timing

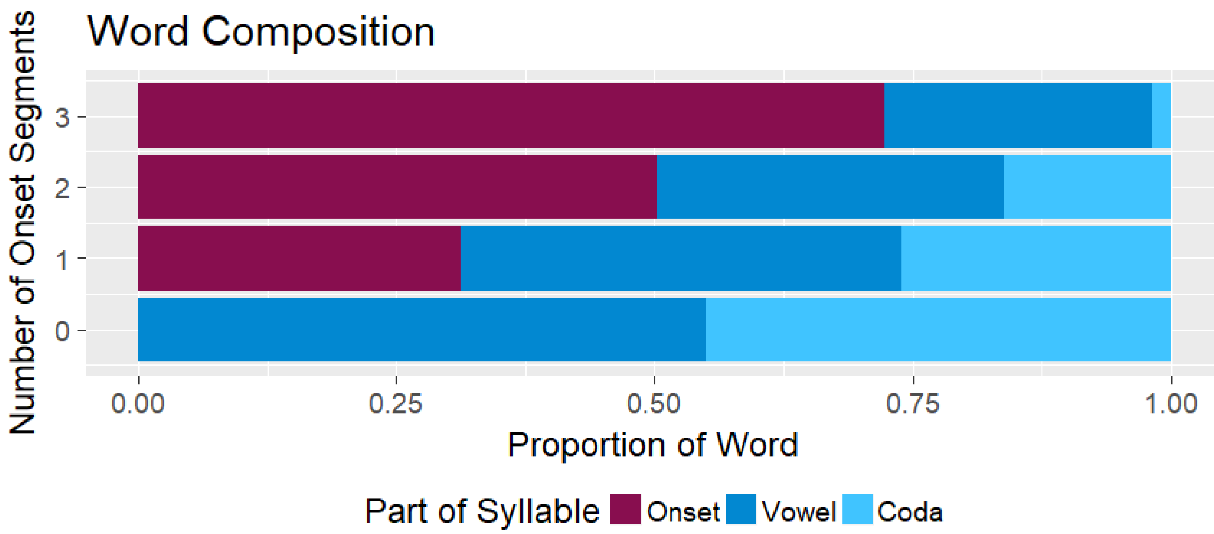

Figure 4 This plot illustrates the average composition of monosyllabic words with zero, one, two, or three onset segments. The boundary between onset (maroon) and vowel (dark blue) shifts to the right as the number of onset segments increases, illustrating that the rime comprises a smaller portion of words with more onset segments. The relatively tiny segment occupied by the coda in words with three onset segments is primarily an artefact of segmentation. Seven of the eight target words which had three onsets segments ended in an oral stop (i.e. straight; see Table 2 in the Appendix for a complete list of target items). Segmentation criteria were such that word-final oral stops had no duration (see §6 for segmentation criteria).

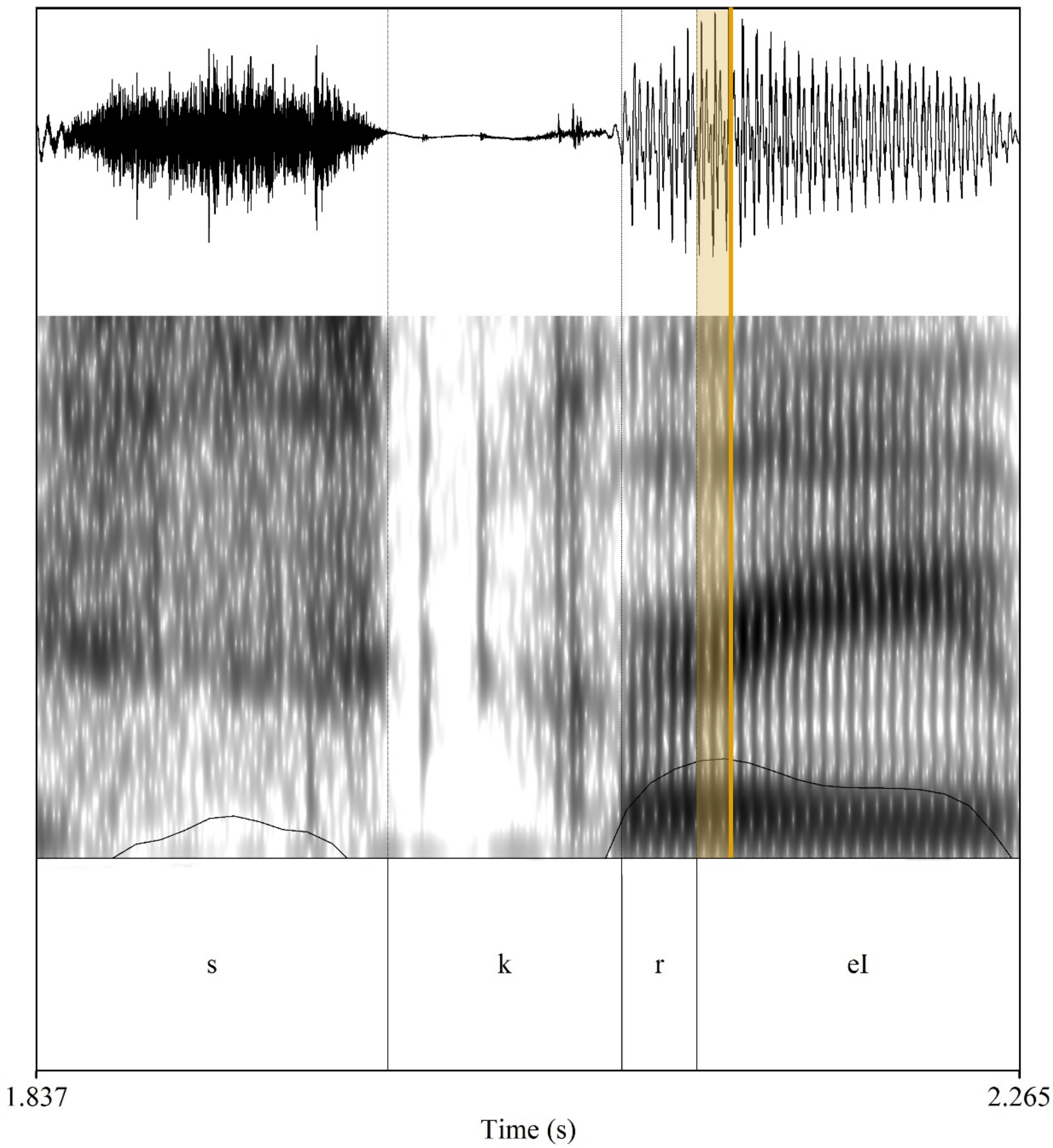

Figure 5 This plot illustrates the calculation of the intensity maximum's timing relative to the onset of the vowel for one token of the word scrape. The intensity contour is plotted as the black trace over the spectrogram and segment boundaries are marked by vertical lines spanning the height of the figure. The dark, orange vertical line marks the intensity maximum, which as expected, falls within the vocalic segment. The semitransparent, lighter orange box spans the distance from the beginning of the vowel to the intensity maximum. The width of this box represents the measure of the timing of the intensity maximum relative to the onset of the vowel. Because the dark orange line occurs later than the semitransparent box, the value of intensity maximum's timing relative to the beginning of the vowel for this token is positive. This measure is analogous to the positive VOT which is calculated when the beginning of vocal fold vibration occurs after the release of a plosive.

of the intensity maximum ($\chi^2 = 523.45, p < .0001$). Post hoc comparisons using the Dunn's test indicate that each onset complexity condition differs significantly from all other conditions at $p < .0001$. While the intensity maximum occurs later during the word as onset complexity increases, it occurs significantly earlier relative to the onset of the vowel. Similar to the timing of the amplitude maximum, the timing of the f0 maximum was calculated relative to the beginning of the vowel, and like the amplitude maximum, it also occurs significantly earlier relative to the onset of the vowel. A Kruskal-Wallis test comparing the effect of onset complexity on the timing of the f0 maximum shows a significant effect of onset complexity on the timing of the f0 maximum ($\chi^2 = 66.79, p < .0001$), and a post hoc Dunn's test indicates that words with zero, one, or two onset segments each differ significantly from one another at $p < .0001$. Words with three onset segments differ significantly from those with one onset segment ($p < .002$) but not from those with two onset segments.

The locations of pitch and amplitude maxima relative to the onset of the vowel correspond well to the pattern of p-center measurements Ryan observed across onset complexities. These results suggest that pitch and amplitude maxima function as adequate phonetic proxies for the p-center, and their correlation with p-center behavior may suggest new approaches to the p-center, as described in the discussion. Goedemans (1998), like also observed that the p-center location was sensitive to the location of the intensity peak. However, Goedemans failed to find evidence that manipulation of the timing of the intensity peak corresponded to changes in the perceived duration of syllables. Goedemans' failure to find perceptual effects tied to the location of the amplitude maxima may have resulted from measurement of the amplitude maximum independent of its relationship to the onset of the vowel. Given Goedemans' results, the location of the amplitude maximum relative to the onset of the vowel likely reflects a more meaningful phonetic effect of onset complexity than does absolute location of the amplitude maximum. Similarly, it may be the case that the location of pitch and amplitude maxima relative to the onset of the vowel together synergistically impact perception of acoustic prominence.

4.3. Maximum change in intensity

Like the intensity maximum and the f0 maximum, the point of maximum change in intensity was also calculated relative to the onset of the vowel. This measurement was calculated by subtracting the time at which the maximum change in intensity occurred from the time at which the vowel segment of the syllable began. This value is negative if the point of maximum change in intensity occurs prior to the beginning of the vowel segment and is positive if the point of maximum change in intensity occurs following the beginning of the vowel segment.

A Kruskal-Wallis test comparing the effect of onset complexity on the timing of the word's maximum change in intensity shows a significant effect of the number of onset segments on the timing of the maximum change in intensity ($\chi^2 = 603.66, p < .0001$). Post hoc comparisons using Dunn's test indicate that each onset complexity condition differs significantly from all other conditions at $p < .001$. As onset complexity increases, both the intensity maximum and the point of maximum change in intensity occur earlier relative to the onset of the vowel. Additionally, as onset complexity increases, the maximum slope of the intensity contour also increases. Performing a Kruskal-Wallis test to compare the effect of onset complexity on the maximum change in intensity, we find the maximum slope of the intensity contour varies significantly by the number of onset segments ($\chi^2 = 155.09, p < .0001$). Corrected post hoc comparisons using Dunn's test indicate that each onset complexity condition differs significantly from all other conditions at $p < .005$ except for the difference between the zero-onset and one-onset conditions ($p = .76$). Figure 7 illustrates these differences in the timing and value of the steepest increase in intensity.

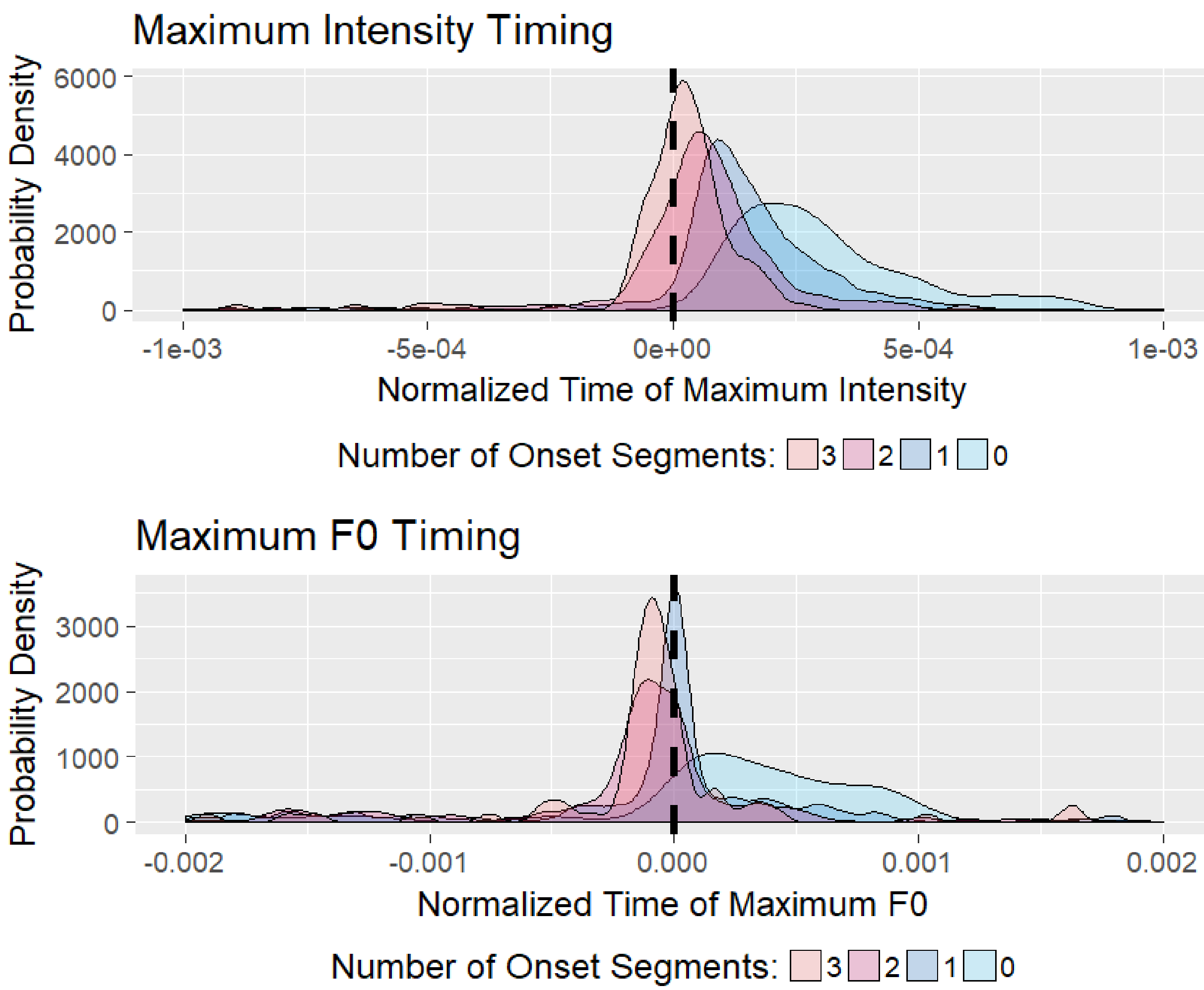

Figure 6 The timing of maximum intensity (top) and the timing of maximum f0 (bottom) are calculated relative to the beginning of the vowel, which is represented by the vertical dashed line.

Together, the timing of the intensity maximum and the timing and value of the maximum change in intensity paint a distinctive picture of the impact of the onset on the shape of the intensity contour. A typical syllable with three onsets may have an intensity maximum right at the beginning of the vowel and a 500 dB/s increase in intensity preceding that maximum, while a typical syllable with only one onset segment may have an intensity maximum occurring further into the vowel and a 300 dB/s increase in intensity occurring right at the beginning of the vowel. A typical syllable with two onsets likely has an intensity contour intermediate to those of syllables with one or three onsets, and a syllable without onsets may have the latest intensity maximum and lowest maximum change in intensity occurring well within the vowel.

4.4. Voicing

The observation that syllables containing phonologically voiceless onsets tend to be heavier than those containing voiced onsets is possibly motivated by differences in f0. Crosslinguistically the

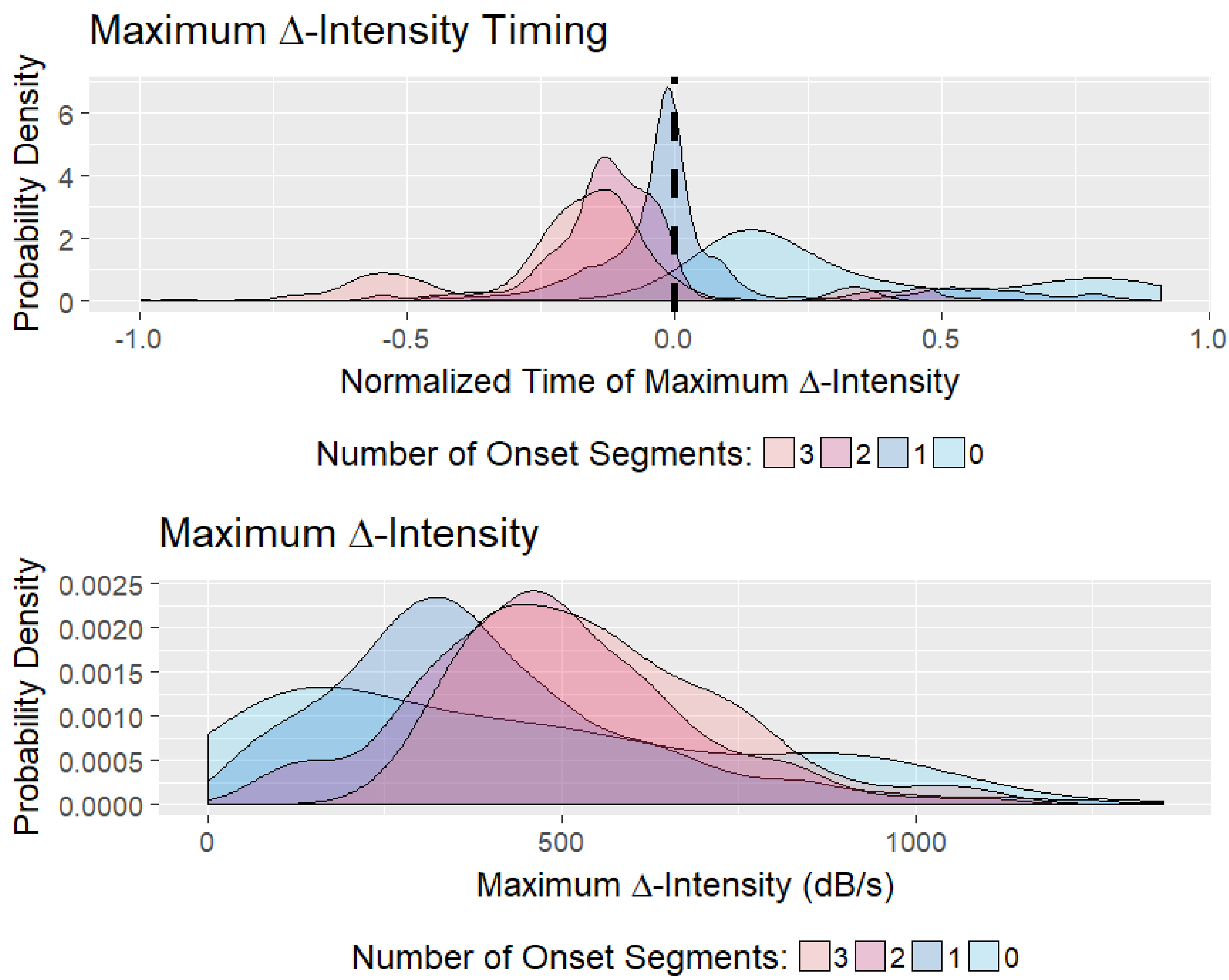

Figure 7 The timing of maximum change in intensity relative to the beginning of the vowel (top) and the value of the slope at that time (bottom).

tendency for high tone to favor voiceless onset segments suggests a functional relationship between voiceless onset segments and higher f0 (Kingston, 2011; Tang, 2008; Yip, 2002). In the literature, this tendency may be most commonly evidenced in cases of tonogenesis where voiceless onsets give rise to the percept of higher pitch which is enhanced and subsequently phonologized as high tone. However, it is also possible that the microtonal perturbation of f0 by voiceless onsets be exploited for syllable weight rather than phonologized as tone, as proposed by Topintzi (2010). Given that high tone preferentially associates with stress (de Lacy, 2002), the marginally higher f0 of syllables containing voiceless onsets may be perceived as a tonal prominence capable of attracting stress, even if the f0 prominence is not phonologized as tone itself. In this way, the f0 associated with voiceless onset segments may contribute to their weight.

Differences in the intensity of voiced and voiceless onset segments may also contribute to their weight. The auditory recovery account predicts that lower intensity onsets will maximally boost the perception of the rime's loudness and thus contribute to its weight. Voiceless segments tend to be of lower intensity than their voiced counterparts due to the absence of vocal fold vibration, and as such, the beginning of the rime is more salient following a voiceless onset segment than following a voiced onset segment. In this way, Gordon (2005) incorporated the greater weight of voiceless onsets in Pirahã into his perceptual account of onset weight effects. If the gradient weight of onsets in English is due to a general auditory mechanism like auditory recovery, we should expect to find an interaction between onset voicing and aspects of the intensity contour in addition to aspects of the f0 contour.

Compounded with differences in intensity, duration may also play a role in the tendency for syllables containing voiceless onsets to be heavier than those containing voiced onsets. In English, voiceless stops are aspirated in single-onset monosyllabic words, increasing their length relative voiced stops in the same context, and voiceless fricatives are generally longer than voiced fricatives as well (Klatt, 1976; Umeda, 1977). Alongside differences in intensity, durational differences may additionally motivate the gradient weight of voiceless onsets in English. By both the auditory recovery account in Gordon (2005) and the p-center account in Ryan (2014), a substantial impact of raw onset duration on syllable weight would be unexpected, but when combined with possible differences in intensity, becomes more plausible. For the auditory recovery account, Ryan (2014) notes that recovery reaches saturation after approximately 40ms (Delgutte, 1982, p. 135), which could account for differences between voiced and voiceless single onset words, but would not be able to account for gradient weight differences across onset complexities. Similarly, for a p-center account of syllable weight, the p-center typically occurs after the bulk of the onset, so raw durational differences in onset consonants alone would not influence weight unless those durational differences themselves were responsible for shifting the p-center earlier in the word.

Analysis of onset voicing is restricted to the set of single onset words, and a number of Kruskal-Wallis tests were conducted to compare the effect of onset segment voicing on normalized onset duration, mean onset intensity, mean word f0, maximum f0, maximum intensity, and the timing of the f0 and intensity maxima. As observed in previous studies including Klatt (1976) and Umeda (1977), a Kruskal-Wallis test comparing the effect of onset segment voicing on normalized duration was significant ($\chi^2 = 251.05$, $p < .0001$; Dunn's post hoc: $p < .0001$). Additionally, a Kruskal-Wallis test conducted to compare the effect of onset segment voicing on the average f0 of the word was significant ($\chi^2 = 5.79$, $p = .02$; Dunn's post hoc: $p < .01$), as was a test to compare the effect of onset segment voicing on the average intensity of the onset ($\chi^2 = 44.08$, $p < .0001$; Dunn's post hoc: $p < .0001$). Words with a voiced onset had lower maximum and average pitch than words with voiceless onsets, and voiced onsets were of greater average intensity than voiceless onsets.

However, onset voicing only impacted the maximum f0 of the overall word ($\chi^2 = 13.5$, $p < .0001$; Dunn's post hoc: $p = .0001$) and was not significant in a Kruskal-Wallis test comparing the effect of onset voicing on maximum intensity of the word ($\chi^2 = .31$, $p = .53$). A Kruskal-Wallis test comparing the effect of onset voicing on the timing of the word's f0 maximum also showed a significant effect of onset voicing ($\chi^2 = 4.98$, $p = .03$), and Dunn's post hoc test indicates that the effect is significant at $p = .013$. Similarly, Kruskal-Wallis test comparing the effect of voicing on the timing of the word's amplitude maxima showed a significant effect of onset voicing ($\chi^2 = 11.33$, $p < .0001$), and Dunn's post hoc test indicates that the effect is at $p < .0005$. Further analysis of the intensity contour indicates that onset voicing is associated with an advance the timing of the maximum increase in intensity ($\chi^2 = 39.35$, $p < .0001$; Dunn's post hoc: $p < .0001$) and an increase in the maximum slope of the intensity contour ($\chi^2 = 63.18$, $p < .0001$; Dunn's post hoc: $p < .0001$). These results suggest all together that the phonetic impact of onset voicing on the syllable is similar to the impact of onset complexity on the syllable, a conclusion which coheres well with the observed participation of onset voicing in syllable weight.

4.5. Summary

The current study finds that onset complexity impacts the acoustic realization of the syllable through its impact on vowel duration and the word's f0 and intensity contours. While peak pitch and intensity occur later in the word as onset complexity increases, they occur earlier relative to the beginning of the vowel. Similarly, the timing and value of the maximum increase in intensity occurs earlier relative to the beginning of the vowel as onset complexity increases. These measures behave similarly to measurements of the p-center collected by Ryan (2014) suggesting that the timing of these maxima and the value of the maximum change in intensity may be capable of acting as acoustic proxies for the p-center when less precise measurements of its location are required. In addition to p-center effects, normalized vowel duration decreases as onset complexity increases, a result previously found in the Buckeye Corpus by Ryan (2014). However, this decrease in vowel duration does not compensate for the overall increase in word duration with onset complexity, suggesting a role for raw syllable duration in onset weight sensitivity.

5. Discussion

The current study robustly replicates results from the literature, showing that the addition of segmental material to the onset impacts vowel duration and syllable f0 and intensity contours. In stimuli controlled for stress and categorical weight, these results reflect general acoustic properties of onset complexity in the language. Because the phonetic results from this general study of onset properties align with those reported for onsets participating in stress assignment, we suggest that the gradient weight afforded syllables due to their onset characteristics is likely universal among languages with weight sensitive stress. Phonetic characteristics of the onset influence aspects of syllable acoustics which are integral to the perception of weight yet exert so small an effect on the overall phonetic realization of the syllable as to evade cooptation into categorical weight systems.

The notion of phonetic effectiveness proposed by Gordon (2002) succinctly captures the reason onsets escape inclusion in categorical weight criteria: their phonetic influence is not great enough to adequately partition the syllable structure inventory in perceptual space. However, onsets still manage to influence weight-based processes like stress assignment in languages like English, provoking questions as to the way in which they accomplish this and the manner of their relationship to categorical weight criteria. Here it is argued that the weight system of English appropriates the acoustic effect that onsets have on the realization of the syllable as a form of prominence enhancement, and the evidence supporting this claim are the data from a production study showing that the acoustic consequences of onsetful syllables mirror the acoustic consequences of categorical weight, independent of their participation in stress-assignment. Under this interpretation of the data, gradient weight criteria are any phonetic factors which mimic the acoustic realization of categorical weight in the language, and gradient weight criteria probabilistically participate in weight-based phenomena such that their perceptual prominence may be consolidated with a phonologically prominent domain, such as a stressed syllable (Ryan, 2014), a metrically strong syllable (Kelly, 2004; Ryan, 2014), or a phrasally stressed syllable (Ryan, 2018).

Some of the results reported in this study raise questions which may be best addressed by investigation of their articulatory basis. Syllables without onsets demonstrate the greatest amount of variability in their acoustic realization across the set of measures taken in this study. This trend can be seen from the broader, flatter contours of the probability densities associated with onsetless syllables compared to any onsetful distribution in Figures 3, 5, and 7. For example, the broader distribution of the timing of the intensity maximum and the value and timing of the maximum change in intensity relative to distributions of onsetful syllable types suggest that onsets place greater constraints on the shape of the syllable's intensity contour than those imposed by the rime alone. This pattern may be attributable to the precision required to correctly coordinate the articulation of consonantal gestures, resulting in the more peaked, less variable distributions of measures taken from onsetful syllables across these measures. Although the difference in acoustic measures across onset complexities is small, these data suggest that acoustic differences caused by onset complexity are relatively reliable. The reliability of these differences may act as an additional factor in the gradient influence of onsets on weight-based processes.

All licit three segment onset clusters of English end in a sonorant, and most licit two segment onset clusters do so as well. In this way, onsets with greater complexity tend to rise in sonority more gradually than those with fewer onset segments. One might expect the more gradual increase in sonority typical of syllables with more complex onsets to accompany a more gradual increase in intensity from the onset to the nucleus, but this was not found to be the case. Rather, greater onset complexity correlated with steeper increases in intensity. The reason for this finding is not clear but may be associated with articulatory correlates of onsetful syllables. As more onset segments are added to the syllable, the timing between the achievement of the rightmost onset consonant's target and the vowel's target decreases, a result attributed to the C-center effect in English (Browman & Goldstein, 1988; Byrd, 1996). The reduced distance between these two targets could result in the steeper maximum increase in intensity observed for syllables with more complex onsets if it accompanies a more rapid decrease of constriction in the oral tract. Investigation of the relationship between the acoustic and articulatory repercussions of onset complexity would be necessary to assess the validity of this speculation.

In addition to exploration of the articulatory basis for acoustic findings reported here, the results of the current study also suggest future avenues for exploration of the p-center phenomenon and its relationship to the perception of loudness. As mentioned in the description of the theory of auditory adaptation, the perception of weight is hypothesized to correlate with the strength of neural firing temporally entrained to the syllable in question. Gordon argues that this measure is best represented acoustically as integrated amplitude over time. The p-center tends to occur prior to the onset of the vowel and the amplitude maximum. The location of maximal change in amplitude is a possible point which meets these temporal characteristics and involves calculation over the amplitude domain. If it is the case that the p-center acts as the point after which amplitude information is integrated for the perception of loudness (and by extension, weight), then the locus of maximal change in intensity within the syllable domain may correspond to that point.

Focus on the slope of change in amplitude would also tie in neatly with the explanation offered by the auditory adaptation account for the primacy of obstruent onset weight over sonorant onset weight. The increase in amplitude from a sonorant onset into a vowel is more gradual than the increase in amplitude from an obstruent onset into a vowel. The shallower slope of the amplitude envelope may allow for greater adaptation of auditory nerve fibers, dampening the rate of neural firing. In this way, incorporation of the measure of amplitude slope into calculation of the p-center and perceptual loudness may paint a more precise picture of the role of amplitude in speech perception.

6. Conclusion

The current paper presents results from a production study of English monosyllabic words showing the impact of onset complexity on acoustic characteristics of the syllable and finds that onset effects in English, a language in which onsets do not participate in categorical weight criteria, nevertheless correspond to effects observed in languages in which they do. Given the similarity of these results to those found in languages with categorical onset weight, these findings suggest that intrinsic phonetic effects of onset complexity are exploited by the phonological weight system to enhance syllable prominence in American English.

7. Acknowledgements

I would like to thank Yolanda Chow, Iliana De Dios, Natalia Reed, Lin Tian, and Joyee Wong for their assistance with data segmentation as well as members of UCSD's PhonCo working group, Eric Baković, Marc Garellek, and Gabriela Caballero for comments and discussion. An earlier version of this project was presented at the Winter 2017 Meeting of the Acoustical Society of America.

Notes

[1]Among them Khalkha Mongolian, a Malayalam-type language in which CVV outweighs CVC and CV yet demonstrates no difference in the length of vowels across open and closed syllables. Gordon suggests Khalkha's mora assignment may differ from that of Malayalam.

[2]Because this test is conducted on normalized onset duration, only onsets containing one, two, or three segments are included. Onsetless target items were excluded from this analysis.

References

Allen, G. (1972). The location of rhythmic stress beats in English: An experimental study i. *Language and speech*, *15*(1), 72–100. doi:https://doi.org/10.1177/002383097201500110

Allen, W. (1973). *Accent and rhythm*. doi:10.2307/370132

Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. In *Proceedings of the institute of phonetic sciences* (Vol. 17, *1193*, pp. 97–110). Amsterdam.

Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer [computer program] version. 5.3. 21. 2012. Retrieved from http://www.fon.hum.uva.nl/praat/

Broselow, E., Chen, S., & Huffman, M. (1997). Syllable weight: Convergence of phonology and phonetics. *Phonology*, *14*(1), 47–82. doi:https://doi.org/10.1017/S095267579700331X

Browman, C., & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, *45*(2-4), 140–155. doi:https://doi.org/10.1159/000261823

Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of phonetics*, *24*(2), 209–244. doi:https://doi.org/10.1006/jpho.1996.0012

Chomsky, N., & Halle, M. (1968). The sound pattern of English.

Cooper, A., Whalen, D., & Fowler, C. (1986). P-centers are unaffected by phonetic categorization. *Perception & Psychophysics*, *39*(3), 187–196. doi:https://doi.org/10.3758/BF03212490

Davis, S. (1988). Syllable onsets as a factor in stress rules. *Phonology*, *5*(1), 1–19. doi:https://doi.org/10.1017/S0952675700002177

Day, R., & Quinn, G. (1989). Comparisons of treatments after an analysis of variance in ecology. *Ecological monographs*, *59*(4), 433–463. doi:https://doi.org/10.2307/1943075

de Lacy, P. (2002). The interaction of tone and stress in Optimality Theory. *Phonology*, *19*(1), 1–32. doi:https://doi.org/10.1017/S0952675702004220

Delgutte, B. (1982). Some correlates of phonetic distinctions at the level of the auditory nerve. *The representation of speech in the peripheral auditory system*, 131–149.

Delgutte, B. (1986). Analysis of French stop consonants with a model of the peripheral auditory system. *Invariance and Variability of Speech Processes. Lawrence Erlbaum Associates, Hillsdale, NJ*, 131–177.

Delgutte, B. (1997). Auditory neural processing of speech. *The handbook of phonetic sciences*, 507–538.

Delgutte, B., & Kiang, N. (1984). Speech coding in the auditory nerve: I. Vowel-like sounds. *The Journal of the Acoustical Society of America*, *75*(3), 866–878. doi:https://doi.org/10.1121/1.390596

Downing, L. (1998). On the prosodic misalignment of onsetless syllables. *Natural Language and Linguistic Theory*, *16*(1), 1–52. doi:https://doi.org/10.1023/A:1005968714712

Duanmu, S. (1994). Against contour tone units. *Linguistic inquiry*, *25*(4), 555–608.

Dunn, O. (1964). Multiple comparisons using rank sums. *Technometrics*, *6*(3), 241–252. doi:https://doi.org/10.1080/00401706.1964.10490181F

Everett, D., & Everett, K. (1984). Syllable onsets and stress placement in Pirahã. In *Proceedings of the West Coast Conference on Formal Linguistics* (Vol. 3, pp. 105–116).

Fowler, C., & Tassinary, L. (1981). Natural measurement criteria for speech: The anisochrony illusion. *Attention and performance IX*, *9*, 521–535.

Goedemans, R. (1998). Weightless segments. *The Hague: Holland Academic Graphics*.

Gordon, M. (2002). A phonetically driven account of syllable weight. *Language*, 51–80. doi:https://doi.org/10.1353/lan.2002.0020

Gordon, M. (2005). A perceptually-driven account of onset-sensitive stress. *Natural Language and Linguistic Theory*, *23*(3), 595–653. doi:https://doi.org/10.1007/s11049-004-8874-9

Gordon, M. (2007). *Syllable weight: Phonetics, phonology, typology*. doi:https://doi.org/10.4324/9780203944028

Hajek, J., & Goedemans, R. (2003). Word-initial geminates and stress in Pattani Malay. *Linguistic review*, *20*(1), 79–94. doi:https://doi.org/10.1515/tlir.2003.003

Ham, W. (2013). *Phonetic and phonological aspects of geminate timing*. doi:https://doi.org/10.4324/9781315023755

Harsin, C. (1997). Perceptual-center modeling is affected by including acoustic rate-of-change modulations. *Perception & psychophysics*, *59*(2), 243–251. doi:https://doi.org/10.3758/BF03211892

Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic inquiry*, *20*(2), 253–306.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. University of Chicago Press.

Hubbard, K. (1994). *Duration in moraic theory* (Doctoral dissertation, University of California at Berkeley).

Hubbard, K. (1995). 'Prenasalised consonants' and syllable timing: Evidence from Runyambo and Luganda. *Phonology*, *12*(2), 235–256. doi:https://doi.org/10.1017/S0952675700002487

Hyde, B. (2007). Issues in Banawá prosody: Onset sensitivity, minimal words, and syllable integrity. *Linguistic Inquiry*, *38*, 239–285. doi:https://doi.org/10.1162/ling.2007.38.2.239

Hyman, L. (1977). On the nature of linguistic stress. *Southern California Occasional Papers in Linguistics*, *Studies in stress and accent*, 37–82.

Hyman, L. (1992). Moraic mismatches in Bantu. *Phonology*, *9*(2), 255–265. doi:https://doi.org/10.1017/S0952675700001603

Hyman, L. (2003). *A theory of phonological weight*. Center for the Study of Language and Information.

Jakobson, R. (1931). *Die betonung und ihre rolle in der wort-und syntagmaphonologie...* Státní tiskárna.

Katz, J. (2010). *Compression effects, perceptual asymmetries, and the grammar of timing* (Doctoral dissertation, Massachusetts Institute of Technology).

Kelly, M. (2004). Word onset patterns and lexical stress in English. *Journal of Memory and Language*, *50*(3), 231–244. doi:https://doi.org/10.1016/j.jml.2003.12.002

Kingston, J. (2011). Tonogenesis. *The Blackwell companion to phonology*, 1–30.
Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *The Journal of the Acoustical Society of America*, *59*(5), 1208–1221. doi:https://doi.org/10.1121/1.380986
Kruskal, W., & Wallis, W. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, *47*(260), 583–621. doi:https://doi.org/10.1080/01621459.1952.10483441
Maddieson, I. (1993). Splitting the mora. *UCLA Working papers in phonetics*, *83*(9), 18.
McCarthy, J., & Prince, A. (1994). The emergence of the unmarked: Optimality in prosodic morphology.
McCarthy, J., & Prince, A. (1995). Faithfulness and reduplicative identity. *Linguistics Department Faculty Publication Series*, 10.
Morton, J., Marcus, S., & Frankish, C. (1976). Perceptual centers (p-centers). *Psychological Review*, *83*(5), 405. doi:https://doi.org/10.1037/0033-295X.83.5.405
Nanni, D. (1977). Stressing words in-ative. *Linguistic Inquiry*, 752–763.
Peirce, J. (2007). Psychopy—psychophysics software in Python. *Journal of neuroscience methods*, *162*(1-2), 8–13. doi:https://doi.org/10.1016/j.jneumeth.2006.11.017
Pompino-Marschall, B. (1989). On the psychoacoustic nature of the p-center phenomenon. *Journal of phonetics*.
Rapp-Holmgren, K. (1971). A study of syllable timing. *Speech Transmission Laboratory–Quarterly status and progress report*, *12*, 14–19.
Ryan, K. (2011). Gradient syllable weight and weight universals in quantitative metrics. *Phonology*, *28*(3), 413–454. doi:https://doi.org/10.1017/S0952675711000212
Ryan, K. (2014). Onsets contribute to syllable weight: Statistical evidence from stress and meter. *Language*, *90*(2), 309–341. doi:https://doi.org/10.1353/lan.2014.0029
Ryan, K. (2018). Prosodic end-weight reflects phrasal stress. *Natural Language & Linguistic Theory*, 1–42. doi:https://doi.org/10.1007/s11049-018-9411-6
Steriade, D. (1991). Moras and other slots. *Proceedings of the Formal Linguistics Society of Midamerica*, *1*, 254–280.
Steriade, D. (1997). Phonetics in phonology: The case of laryngeal neutralization.
Steriade, D. (1999). Alternatives to syllable–based accounts of consonantal phonotactics. *Proceedings of the LP*, 205–245.
Strehlow, T. (1942). Aranda phonetics. *Oceania*, *12*(3), 255–302. doi:https://doi.org/10.1002/j.1834-4461.1942.tb00360.x
Tang, K. (2008). *The phonology and phonetics of consonant-tone interaction* (Doctoral dissertation).
Topintzi, N. (2010). *Onsets: Suprasegmental and prosodic behaviour*. doi:https://doi.org/10.1017/CBO9780511750700
Topintzi, N., & Nevins, A. (2017). Moraic onsets in Arrernte. *Phonology*, *34*(3), 615–650. doi:https://doi.org/10.1017/S0952675717000306
Trubetzkoy, N. (1939). Grundzüge der phonologie. Göttingen: Van der Hoeck and Ruprecht. translated 1969 by Christine Baltaxe as Principles of Phonology. Berkeley and Los Angeles: University of California Press.
Umeda, N. (1977). Consonant duration in American English. *The Journal of the Acoustical Society of America*, *61*(3), 846–858. doi:https://doi.org/10.1121/1.381374
Villing, R. (2010). *Hearing the moment: Measures and models of the perceptual centre* (Doctoral dissertation, National University of Ireland Maynooth).
Yip, M. (2002). *Tone*. doi:https://doi.org/10.1017/CBO9781139164559

Zec, D. (1988). *Sonority constraints on prosodic structure* (Doctoral dissertation, Stanford University).
Zhang, J. (2001). *The effects of duration and sonority on contour tone distribution–typological survey and formal analysis* (Doctoral dissertation).

Table 2 Target Items. Listed by number of onset segments.

0	1		2		3
	rap		trap		scrap
	cap		crap		strap
	sap		snap		
	chap				
ink	link		slink		
	rude		crude		screwed
ape	gape		grape		
	cape		crepe		scrape
ate	fate	gate	great	slate	straight
	rate	Kate	crate	freight	
	late	Tait	prate		
	date		plate		
ale	nail	dale	snail		
	rail	gale	trail		
	tale	bail	scale		
	kale	pail	braille		
	sale				
	rot	lot	brought		
	bought	pot	fraught		
	fought	taught	plot		
	cot	dot	Scott		
	sought	fraught	trot		
oath	both				
ash	lash	sash	trash	smash	splash
	rash	gash	crash	stash	
	cash	mash			
ill	hill	till	shrill		
	rill	dill	grill		
	shill	bill	krill		
	kill	pill	trill		
	gill				
	rain	Jane	train		
	sane	bane	drain		
	pain	chain	stain		
oak	soak		stoke		stroke
	poke		smoke		
isle	mile	pile	smile		
	bile				
	rug		plug		
	pug		smug		
	mug		shrug		
	lug				
	rim		trim		
	Tim		grim		
	dim				
	lip	sip	trip		strip
	rip	chip	slip		
	tip	dip			
aim	lame		flame		
	fame		blame		
			frame		
	cough		scoff		
ache	rake	lake	break	flake	
	take	bake	steak	snake	
	sake	Jake	drake		
	dumb		drum		
	rum				

Table 3 Filler items.

floral	baker	mischief	candied	walnut	paper
raisin	lambast	liquid	squirrel	swivel	impound
darling	decent	yellow	sour	splendor	terror
shower	trauma	mucus	mercy	brandy	browser
sheriff	polyp	abyss	valid	vanquish	promise
trifle	market	splinter	awning	pleasure	jury
nausea	limpid	zenith	valor		