

# 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](mailto: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. 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  $f_0$  were collected. Results of the current study are consistent with the predictions of Gordon's perceptual account of categorical weight, showing that integrated intensity of the rime is incapable of driving onset weight behavior in English. Furthermore, results indicate that onsets impact the shape of the intensity envelope in a manner consistent with explanations for gradient onset weight that appeal to onset influence on the perceptual center (Ryan, 2014). Together, these results show that cues to gradient weight act independently of primary cues to categorical weight to probabilistically impact weight sensitive stress assignment in English.

**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; 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 (gradient) 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, Marcus, and Frankish 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 section 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 section 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 called 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,<sup>1</sup> 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, intensity, 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 intensity 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 intensity 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 /<sup>h</sup>məˌveɪtɪv/ vs. manipulative /məˈnɪpjələtɪv/. However, beyond this subset of the English lexicon, categorical onset criteria are not evidenced in the weight systems investigated. 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 (Kelly, 2004; Ryan, 2014).

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 intensity 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. In Gordon's perceptual energy account, auditory recovery during the onset provides a small perceptual boost to the integrated intensity of the rime. Crucially, for the perceptual energy account to predict the typological rarity of categorical onset weight, integrated intensity of the rime must be the primary factor that drives weight criteria. However, if the proportion of the syllable occupied by the rime decreases as onset complexity increases, it becomes more difficult to see how greater onset complexity contributes to greater overall perceptual energy, unless shorter rimes have much greater intensity than longer rimes or the perceptual boost afforded by long onsets is quite strong. In light of these complications for the perceptual energy account (among others described in Ryan 2014, pp. 332), Ryan pursues an alternative perceptual account of syllable weight 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 is influenced by the *perceptual center* (hereon p-center) of the syllable, a concept developed by Morton et al. 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.

Ryan (2014) affirms that multiple phonetic factors likely interact with the p-center to produce gradient weight behavior, including auditory recovery (Gordon, 2005), the syllable’s intensity contour (Goedemans, 1998), and tonal perturbation—though the role that this last factor, tonal perturbation, plays in accounts of syllable weight is much less clear. Differences in  $f_0$  have not definitively been linked to shifts in the p-center (though see Janker and Pompino-Marschall 1991), nor is  $f_0$  used to define weight categories crosslinguistically (Gordon, 2002). However, as both Ryan (2014) and Goedemans (1998) point out,  $f_0$  plays an important role in the acoustic realization of weight. Ryan notes that onset voicelessness covaries with higher  $f_0$  (Kingston, 2011; Tang, 2008; Yip, 2002) and that higher  $f_0$  in turn correlates with stress (de Lacy, 2002), and Goedemans notes that in particular,  $f_0$  is one of the most reliable acoustic cues to stress in English (Fry, 1955, 1958, 1965; Van Heuven & Sluijter, 1996). So, even though  $f_0$  does not play a role in categorical weight criteria, it may still play a role in gradient weight behavior. Specifically, if  $f_0$  contours of onsetful syllables share properties in common with the acoustic realization of weight-based properties like stress, it may motivate their heavy behavior in domains like metered verse. For this reason, the current study treats properties of the  $f_0$  contour as possible acoustic correlates of gradient onset weight in addition to the duration- and intensity-related measures motivated by the p-center and perceptual energy accounts of syllable weight.

### 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 clearly 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 categorical 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 onset complexity impacts syllable acoustics (i.e. duration, intensity, and  $f_0$ ) independently of propensity to receive stress. Onsets in English impact syllable acoustics in a qualitatively similar manner as onsets in languages with *categorical* onset weight and in a manner consistent with phonetic accounts of categorical *rime* weight in English.

In the current study, English speakers pronounced one-syllable words of varying onset complexity and measurements of intensity (peak, average, & integrated), duration (normed onset,

normed vowel & total), and f0 (peak & average) were taken. Measures of average and integrated intensity and duration are motivated by Gordon’s perceptual account of categorical onset weight, and measures of the timing of peak intensity and f0 are motivated by Ryan’s perceptual account of gradient onset weight. This study assesses the extent to which the acoustic factors implicated in each of these accounts is capable of explaining the gradient weight behavior of onsets in English. If acoustic measures associated with categorical weight also play a role in the gradient weight of onsets in English, we expect one of two possibilities: Either syllables with greater onset complexity should have rimes with greater integrated intensity, or the integrated intensity of the rime should remain constant while onset duration increases to allow the perceptual boost afforded by auditory recovery to influence the syllable’s perceived weight. If acoustic measures associated with the p-center play a role in the gradient weight of onsets, we expect peak intensity or f0 to occur earlier in the syllable as onset complexity increases. If both accounts play a role in gradient onset weight, we would expect to see a combination of the results predicted by each account individually. If gradient onset weight is not motivated by general acoustic properties of onsets, we would expect to see either null results or results contrary to the predictions of both accounts.

### 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. One participant did not finish the study, and another did not say the target words within a carrier phrase. Data from these two participants (1 female) were excluded from analysis, for a total of 17 participants reported.

### 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*, *state*, *straight*). Target words fell into one of 8 vowel categories and one of 20 rime categories, as shown in Table 1 (right). 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 rightmost 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 (/ou/, /aɪ/, /eɪ/), and central schwa. The front vowels are

Onset Types			Rime Types		
# Segs.	Onsets	Examples	Vowel	Rime	Examples
0	∅	ate, ash	æ	/æp/	rap, trap
1	/b/, /tʃ/, /d/,	bail, chain, dip	ɪ	/æf/	ash, rash
				/ɪl/	ill, shill
	/f/, /g/, /dʒ/,	fate, gate, Jane	u	/ɪm/	rim, trim
				/ɪp/	rip, trip
	/k/, /l/, /m/,	kale, lash, mash	ə	/ud/	rude, crude
				/əm/	rum, drum
/n/, /p/, /ɹ/,	nail, pill, rip	/əg/	rug, shrug		
/s/, /ʃ/, /t/	sate, shill, tip	ɑ	/ɑt/	ought, taught	
2	/tʃɹ/, /kɹ/, /gɹ/,	train, krill, grape	ou	/ɑf/	cough, scoff
				/ouθ/	oath, both
	/pɹ/, /fɹ/, /bɹ/,	prate, frame, break	aɪ	/ouk/	oak, soak
				/aɪl/	isle, mile
	/ʃɹ/, /dʒɹ/, /pl/	shrill, drake, plot	eɪ	/eɪp/	ape, cape
/sl/, /sn/, /sk/,	slip, snail, scape	/eɪt/	ate, Kate		
/fl/, /bl/	flake, blame	/eɪl/	ale, kale		
3	/skɹ/, /stɹ/, /spl/	scrape, straight, splash	/eɪm/	aim, fame	
			/eɪn/	rain, train	
			/eɪk/	ache, rake	

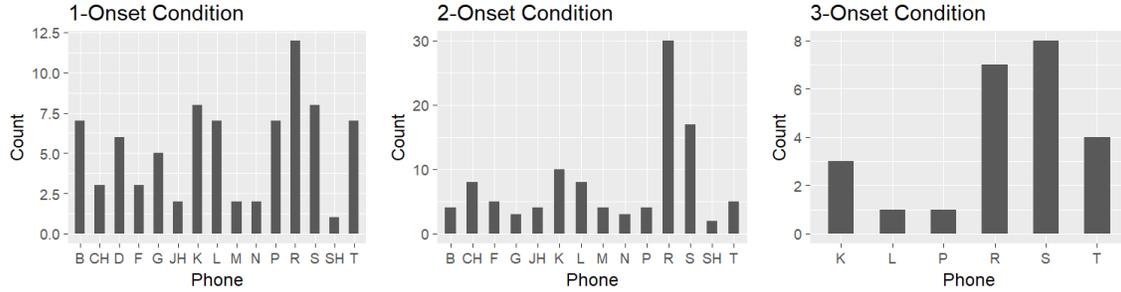
**Table 1** Target items. A representative subset of target items for each onset type (left) and for each vowel and rime type (right) used in the study.

contained in five rime types, the back vowels and schwa account for six rime types, and diphthongs account for nine rime types. Additionally, eleven different coda types are represented across the nineteen rime types. Eight out of the nineteen rime types contain oral stop codas, and nine out of the nineteen rime types have voiced codas. Ten 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, and Table 1 summarizes the set of unique onsets used for each of the four onset complexity conditions. Figure 1 shows that 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.



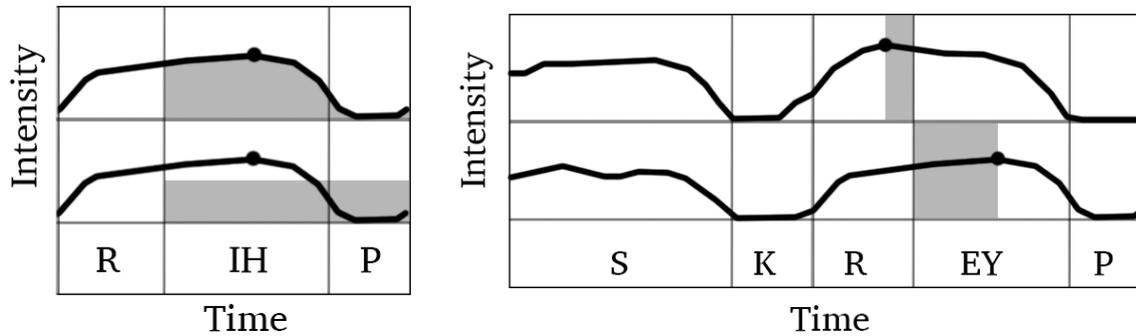
**Figure 1** Target stimuli onset phone frequency across onset complexity conditions. Phones are labeled using ARPABET phonetic transcription codes.

#### 4. Analysis and results

Target words were segmented and labeled using Praat TextGrids. Segmentation was determined using visual landmarks in the spectrogram. For onsetless target words, vowels were segmented from the beginning of irregular voicing following the word *say* to the offset of the target vowel’s second formant. Irregular voicing in this position is due to glottalization associated with the stressed word-initial vowel in a hiatus position (Davidson & Erker, 2014; Garellek, 2013). Otherwise, 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. For stop-initial target words, if the duration between the end of *say* and the stop release burst was less than 300ms and no breath could be heard, word-initial stops were segmented from the offset of the second formant of the vowel in *say* to the end of the stop release burst. For stop-initial target words with greater than 300ms between the end of *say* and the stop release burst, if no breath could be heard, word-initial stops were segmented from the beginning to the end of the stop release burst. If a breath could be heard, word-initial stops were segmented from the end of the breath to the end of the stop release burst. Word-final oral stops were segmented from the end of the vowel to the end of the stop release burst. If glottalized, word-final voiceless stops were segmented from the beginning of irregular voicing to the end of the stop’s release burst, or to the end of irregular voicing if no stop release was present, following (Seyfarth & Garellek, 2015). 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 rime 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).

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



**Figure 2** This plot illustrates the calculation of normalized integrated intensity for one hypothetical token of the word *rip* (left) and the timing of the intensity maximum relative to the beginning of the vowel for two hypothetical tokens of the word *scrape* (right). Each token’s intensity contour is plotted as a black trace and segment boundaries are marked by vertical lines spanning the height of each subfigure. The black dot on each intensity contour marks its maximum. **Left:** The shaded portion underneath the top contour represents the rime’s raw integrated intensity. The shaded rectangle under the bottom contour represents the average intensity multiplied by the rime duration. Normalized integrated intensity is calculated by dividing raw integrated intensity (top) by average integrated intensity (bottom). **Right:** For each token of *scrape*, a gray box spans the distance from the beginning of the vowel to the intensity maximum, and its width represents the value of the intensity maximum’s time relative to the beginning of the vowel. The relative time of the top token’s intensity maximum is negative because the intensity maximum precedes the vowel, and the relative time of the bottom token’s intensity maximum is positive because its maximum occurs during the vowel.

intensity (dB) was calculated for each word, and average intensity (dB) was calculated for the rime of each word. Integrated intensity (dBs) was also calculated for the rime of each word using `scipy.integrate.simps` from the Python library `scipy` (Jones, Oliphant, Peterson, et al., 2001), which performs numerical integration over discrete samples using Simpson’s Rule. To account for variation in speaking level across tokens, the rime’s average intensity was multiplied by the rime duration to create a measure of mean integrated intensity for each token, and integrated intensity was normalized through division by this mean. This creates a normalized measure analogous to Gordon’s (2002) measure of “total energy.” Figure 2 (left) illustrates how this measure was calculated. Additionally, for each word, the timing of the maximum intensity was calculated relative to the beginning of the word by dividing the time point at which the peak intensity 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. Then, 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. This measure is negative if the intensity maximum occurs before the beginning of the vowel and positive if the intensity maximum occurs during the vowel.

The intensity maximum was calculated relative to the beginning of the vowel rather than the beginning of the word because the vowel is most likely to contain the intensity maximum. Of all segments in the word, the vowel is the most sonorous, 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 2 (right) provides a schema representing how this measure was calculated. In the following sections, any reference to intensity maximum timing will refer to this measure.

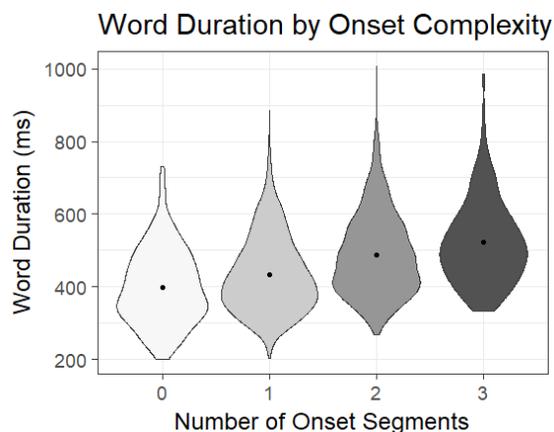
In addition to measures of maximum intensity and its timing, the rate of change in intensity over the course of the word was also calculated. To do so, each sample of the intensity contour and its corresponding time point were logged for each word. These values were then treated as a time series from 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 section 4.4.

An f0 contour was also extracted from each word using Praat's autocorrelation algorithm to detect acoustic periodicity (Boersma, 1993). To set the floor and ceiling parameters adequately for f0 estimation, speakers were impressionistically binned as having 'low' or 'high' voices. A floor of 75Hz and ceiling of 300Hz were used for 'low' voices, a 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 f0 (Hz) was calculated for every word, and average f0 (Hz) was calculated for the nucleus of each word. Additionally, for each word, the timing of the peak f0, and the timing of the maximum change in f0 were calculated relative to the beginning of the vowel in the same manner as described for the intensity contour.

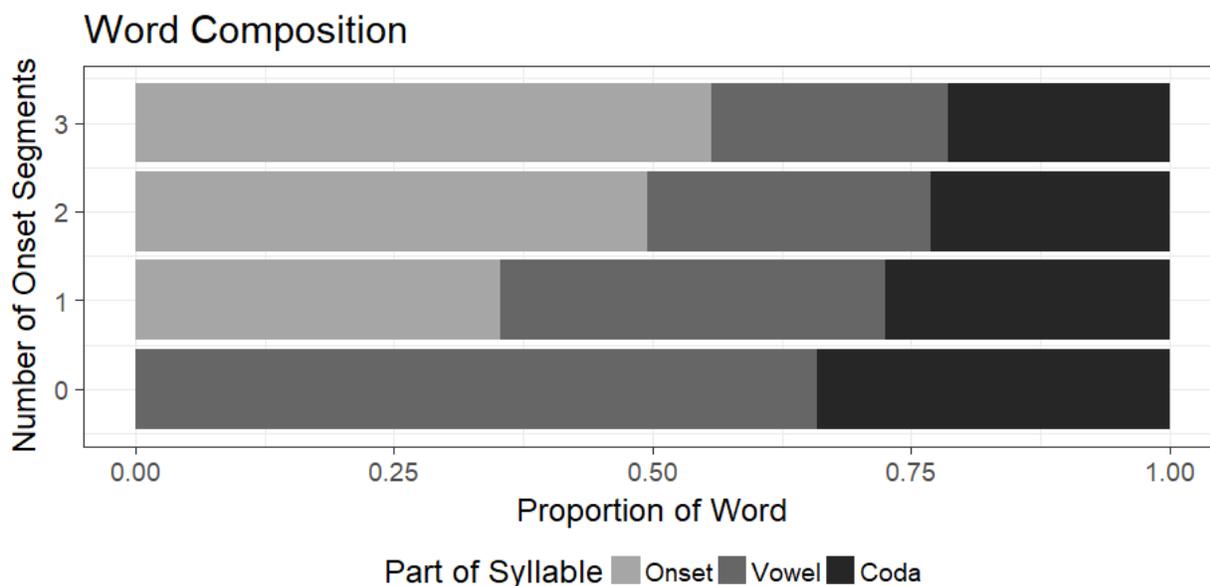
Throughout the analysis, the statistical programming language R (R Core Team, 2017) and the package *lme4* (Bates, Mächler, Bolker, & Walker, 2015) were used to perform linear mixed effects analysis. Each dependent variable discussed in Section 4.1–4.4 was modeled with onset complexity (cond: Levels 0, 1, 2, 3) as the fixed effect and onset, nucleus, coda, rime, and subject identity as random intercept effects ( $\sim \text{cond} + (1|\text{onset}) + (1|\text{nuc}) + (1|\text{coda}) + (1|\text{rime}) + (1|\text{subj})$ ). For each model, normality of residuals was assessed visually, and if the distribution of residuals deviated substantially from normality, the response variable was log-transformed to be compliant with the normality assumption for linear mixed models. Following transformation (if applicable), model residuals did not exhibit obvious deviations from homoscedasticity and (log-) normality. To assess the significance of onset complexity as a predictor for each response variable, likelihood ratio tests were conducted, comparing the fit of the full model against a model without the fixed effect of onset complexity. To provide interpretable effect size estimates for models that were constructed with log transformed data, the difference in effect between three-onset words and zero-onset words is back-transformed into the response variable's original units. The results of each model are reported in more detail below.

#### 4.1. Duration

A linear mixed model was constructed to assess the effect of onset complexity on word duration. The model predicted log-transformed word duration using onset complexity as a fixed effect and onset, nucleus, coda, rime, and subject identity as random effects. Onset complexity affected word duration ( $\chi^2(1) = 36.83$   $p < .0001$ ) such that words with three onset segments were on average 145.89ms longer than words with no onset segments. Figure 3 shows this result. 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



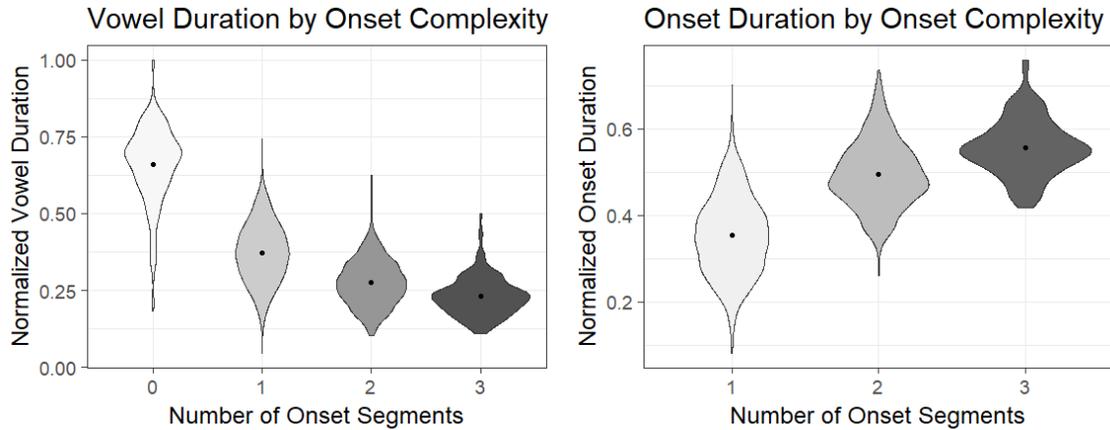
**Figure 3** Raw word duration increases with the number of segments in the syllable onset. Black dots indicate mean values.



**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.

specific to onsets, the result shown in Figure 3 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.

Normalized onset duration was modeled like raw duration using onset complexity as a fixed effect and onset, nucleus, coda, rime, and subject identity as random effects. A likelihood ratio test comparing the fit of the model with onset complexity as a fixed effect versus a model without any fixed effects indicated that onset complexity has a significant effect on the proportion of the



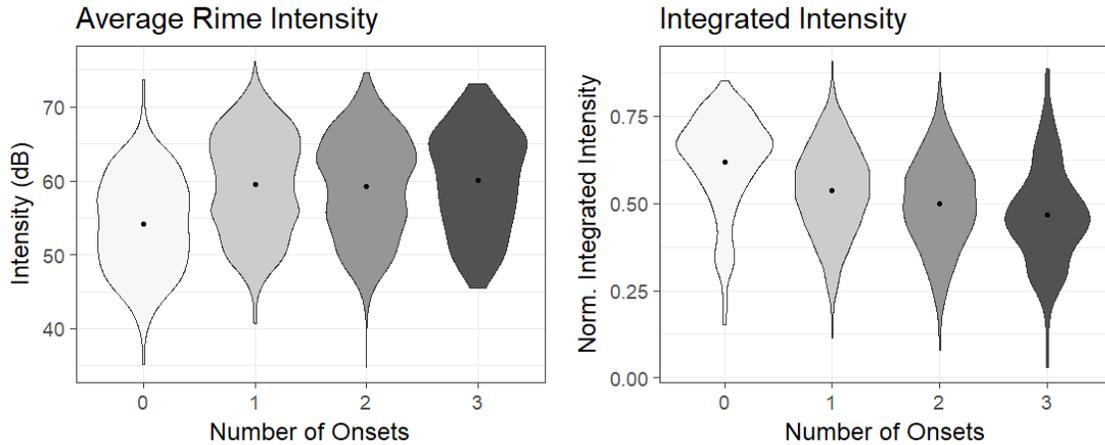
**Figure 5** The proportion of the syllable occupied by the onset is greater in words with more onset segments (right), while the proportion of the syllable occupied by the vowel is lesser in words with more onset segments (left). The normalized duration values plotted on the y-axis of each figure correspond to the duration of the onset (right) or vowel (left) divided by the total word duration, respectively. Black dots indicate value means.

syllable occupied by the onset ( $\chi^2(1) = 48.79, p < .0001$ ). The percentage of the syllable occupied by the onset increased by  $11.38\% \pm 1.13\%$  with each additional onset segment. All together, these results show that as onset complexity increases, so does the proportion of the syllable occupied by the onset (Fig. 5).<sup>2</sup> Conversely, for the log-transformed proportion of the syllable occupied by the vowel, a likelihood ratio test comparing a linear mixed model with a fixed effect of onset complexity against a model without fixed effects indicates that an increase in onset complexity corresponds to a significant decrease in the normalized duration of the vowel ( $\chi^2(1) = 42.94, p < .0001$ ). The vowel occupied on average 26.99% less of the syllable in three-onset words than in zero-onset words. These effects are summarized in Figures 4 & 5.

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 rime is the arbiter of weight both perceptually as well as phonologically, a proportionally smaller rime 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 for acoustic correlates other than rime duration in the assessment of onset contribution to syllable weight. Data from the current study rule out the possibility that absolute rime duration alone drives the influence of onsets on closed-syllable weight.

#### 4.2. Intensity

A likelihood ratio test was conducted to compare a linear mixed model that predicted average rime intensity using onset complexity as a fixed effect and onset, nucleus, coda, rime, and subject



**Figure 6** Raw average intensity increases with onset complexity (left). Normalized integrated intensity decreases with the number of segments in the syllable onset (right). Black dots indicate value means.

identity as random effects against a model that did not include onset complexity as a fixed effect. The test indicated that onset complexity has a significant impact on the average intensity of the rime ( $\chi^2(1) = 9.72, p = .002$ ), such that the average intensity of the rime increases  $.8 \pm .24$  dB with each segment that is added to onset. However, despite the increase in average intensity with added onset complexity, intensity integrated over the rime decreased as onset segments increased. A likelihood ratio test comparing a model with onset complexity as a fixed effect against a model without an onset complexity as an effect found that onset complexity has a significant impact on normalized integrated intensity ( $\chi^2(1) = 23.95, p < .0001$ ), decreasing it by  $.04 \pm .006$  units with each segment added to the syllable onset.

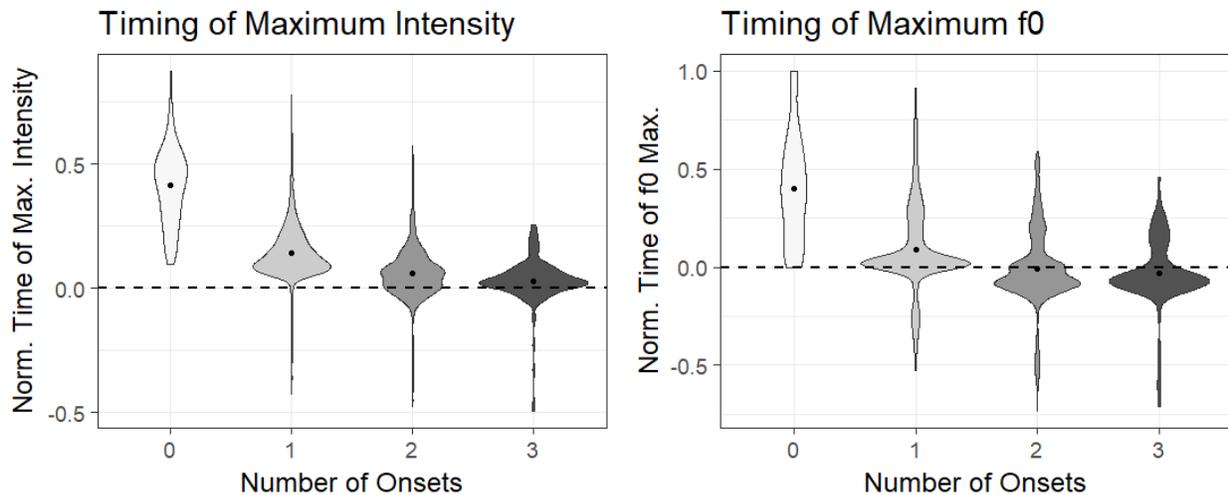
In Gordon's (2002) perceptual account of syllable weight, perceptual energy is derived from total energy (here, integrated intensity) using estimates of perceived loudness taken from Warren (1970). Crucially, perceptual energy is founded on integrated intensity rather than average intensity because psychoacoustic work like Warren's among others (see Moore 2012 for an overview) shows that the ear integrates intensity over time to construct the percept of loudness. Although average rime intensity increases with onset complexity, as shown in Figure 6 (left), integrated intensity decreases (right). Since average intensity increases, the decrease in integrated intensity must be driven by the reduction in rime duration reported in section 4.1. Rime intensity does not increase sufficiently with onset complexity to compensate for the reduction in rime duration, resulting in lower integrated intensity values at higher onset complexities. Together, these results suggest that the impact of onset complexity on the integrated intensity of the rime cannot account for the gradient weight of onsets in English.

#### 4.3. *Timing of intensity and f0 maxima*

A likelihood ratio test comparing a model that includes onset complexity against a model that doesn't indicates that onset complexity plays a significant explanatory role in the timing of the word's intensity maximum ( $\chi^2(1) = 36.17, p < .0001$ ). While the intensity maximum occurs later during the word as onset complexity increases, it occurs significantly earlier relative to the beginning of the vowel. Similar to the timing of the intensity maximum, the timing of the f0 maximum was calculated relative to the beginning of the vowel, and like the intensity maximum,

it also occurs significantly earlier relative to the beginning of the vowel. To assess the impact of onset complexity on the timing of the f0 maximum, the times of the f0 maximum were log-transformed, and two linear mixed models were compared using a likelihood ratio test: One model included onset complexity as a fixed effect, and the other did not. The likelihood ratio test indicated that onset complexity has a significant impact on the timing of the f0 maximum ( $\chi^2(1) = 28.17, p < .0001$ ), such that in words with three onset segments, the f0 maximum occurs on average 27.3% earlier in the word relative to the beginning of the vowel than it does in words with zero onset segments.

The locations of f0 and intensity maxima relative to the beginning of the vowel correspond well to the pattern of p-center measurements Ryan observed across onset complexities. These results suggest that f0 and intensity 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) also observed that the timing of the p-center 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 intensity maxima may have resulted from measurement of the intensity maximum independent of its relationship to the beginning of the vowel. Given Goedemans' results, the location of the intensity maximum relative to the beginning of the vowel likely reflects a more meaningful phonetic effect of onset complexity than does absolute location of the intensity maximum. Similarly, it may be the case that the location of f0 and intensity maxima relative to the beginning of the vowel together synergistically impact perception of acoustic prominence.



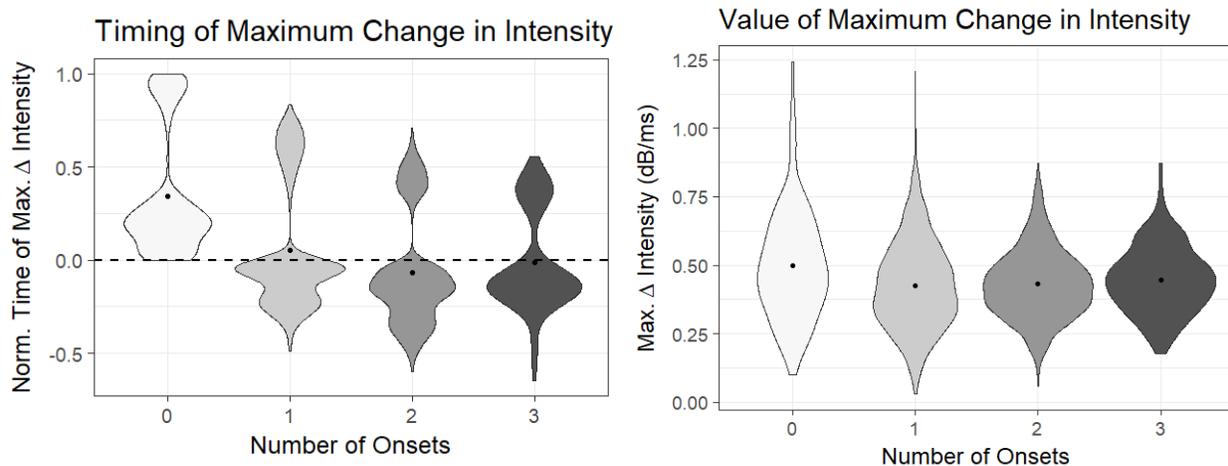
**Figure 7** The timing of maximum intensity (left) and the timing of maximum f0 (right) are calculated relative to the beginning of the vowel, which is represented by the horizontal dashed line.

#### 4.4. *Maximum change in intensity and maximum change in f0*

Like the intensity maximum and the f0 maximum, the point of maximum change in intensity and the point of maximum change in f0 were also calculated relative to the beginning of the vowel. These measurements were calculated by subtracting the normalized time at which the maximum change in intensity or f0 occurred from the normalized time at which the vowel segment of the

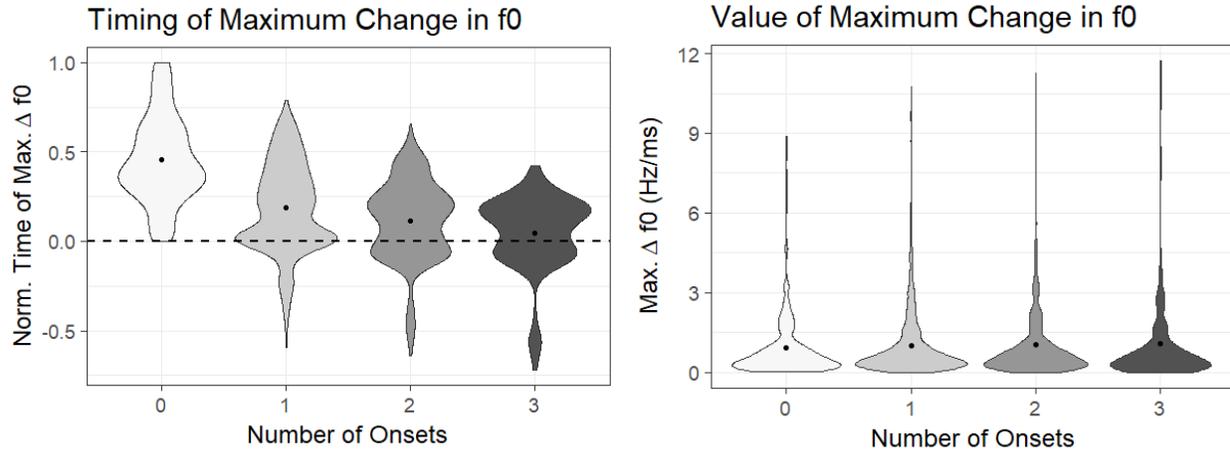
syllable began. These values are negative if the point of maximum change occurs prior to the beginning of the vowel segment and are positive if the point of maximum change occurs following the beginning of the vowel segment.

Two linear mixed models were constructed to assess the effect of onset complexity on the timing of the maximum change in intensity. One modeled the timing of the maximum change in intensity using onset complexity as a fixed effect and onset, nucleus, coda, rime, and subject identity as random effects. The other used only random effects to predict the timing of the maximum change in intensity. A likelihood ratio test comparing the two models indicated that onset complexity has a significant impact on the timing of the word's maximum change in intensity ( $\chi^2 = 19.32(1), p < .0001$ ). With each segment added to the onset, the point of maximum change in intensity occurs  $11.73\% \pm 2.35\%$  earlier relative to the beginning of the vowel. Further, as onset complexity increases, the maximum change in intensity remains the same. A linear mixed model including onset complexity as a fixed effect and a model excluding onset complexity were compared using a likelihood ratio test and were not found to be significantly different from one another ( $\chi^2(1) = 1.13, p = .29$ ). Figure 8 illustrates these differences in the timing and value of the steepest increase in intensity.



**Figure 8** The timing of maximum change in intensity relative to the beginning of the vowel (left) and the value of the slope at that time (right).

Onset complexity impacts the timing and value of the maximum change in  $f_0$  much like it impacts the maximum change in intensity. Two linear mixed models were constructed to assess whether onset complexity explains differences in the timing of the maximum change in  $f_0$  across syllables. One model included onset complexity as a fixed effect, and the other did not. The models were compared using a likelihood ratio test, and were found to be significantly different from one another ( $\chi^2(1) = 29.06, p < .0001$ ). The onset complexity model indicates that with each segment added to the onset, the time point of maximum change in  $f_0$  occurs  $8.93\% \pm 1.31\%$  earlier in the word, relative to the beginning of the vowel. Further, the maximum change in  $f_0$  remains constant across onset complexities. Two linear mixed models were constructed to assess the impact of onset complexity on the log-scaled value of the maximum change in  $f_0$  and were compared using a likelihood ratio test. The test indicated that the model that included onset complexity as a fixed effect was not significantly different from the model that did not include onset complexity ( $\chi^2(1) = 0.3, p = .59$ ). Regardless of the number of segments in the syllable onset,  $f_0$  increases on average by at most  $.97$  Hz/ms (CI = [.92,1.02]). These results are shown in



**Figure 9** The timing of maximum change in f0 relative to the beginning of the vowel (left) and the value of the slope at that time (right).

Figure 9.

Together, the timing of intensity and f0 maxima and the timing and value of their maximum change paint a distinctive picture of the impact of onset complexity on the shape of the intensity and f0 contours. A typical syllable with three onsets may have intensity and f0 maxima right at the beginning of the vowel, while a typical syllable with only one onset segment will likely have intensity and f0 maxima occurring further into the vowel. Across all onset complexities, intensity could increase by as much as  $.4 \pm .02$  dB/ms preceding the intensity maximum and f0 by as much as  $.97$  Hz/ms (CI = [.92,1.02]). These differences in intensity and f0 contours across different levels of onset complexity are qualitatively similar, raising the possibility that they are coordinated in some way— either by shared articulatory constraints or by a shared acoustic goal.

4.5. *Summary*

The current study finds that onset complexity impacts the acoustic realization of the syllable through its impact on rime duration, rime intensity, and the word’s f0 and intensity contours. While peak f0 and intensity occur later in the word as onset complexity increases, they occur earlier relative to the beginning of the vowel. Similarly, the timing of the maximum increase in intensity and f0 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 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 possible role for raw syllable duration in onset weight sensitivity. Although average rime intensity was observed to increase with the addition of segmental material to the onset, integrated intensity of the rime decreases, suggesting that onset weight in English is not likely motivated by the energy of the rime.

5. **Discussion**

The current study robustly replicates results from the literature showing that the addition of segmental material to the onset reduces vowel duration (Browman & Goldstein, 1988; Goedemans, 1998; Ryan, 2014, among others). In addition to these duration results, reduction in the rime's integrated intensity at greater onset complexities casts doubt on the ability of Gordon's (2002, 2005) theory of perceptual weight to account for the gradient weight of onsets in English. However, this conclusion is not necessarily a negative outcome for Gordon's account, since it was designed to predict categorical weight behavior rather than gradient behavior. Gordon's account correctly predicts that categorical weight criteria that make reference to onset complexity are rare. While integrated intensity undoubtedly plays a primary role in the determination of categorical weight, it may be the case that the primary factors motivating gradient weight behavior have a negligible impact on categorical weight, and/or *vice versa*, that the primary factors motivating categorical weight behavior have a negligible impact on gradient weight. More phonetic work comparing correlates of gradient weight with those of categorical weight would be necessary to adjudicate on this point. In any case, the fact that gradient onset weight in English accompanies shorter rimes with lower integrated intensity suggests that acoustic correlates of categorical and gradient weight likely operate somewhat independently of one another.

Predictions made by the p-center account better explain the acoustic impact of onset complexity on the syllable. This study shows that greater onset complexity brings intensity and f<sub>0</sub> maxima earlier in the syllable relative to the beginning of the vowel. Prior to this work, it had been shown that the timing of the intensity maximum covaries with the p-center across syllables of the same onset complexity (de Jong, 1994; Pompino-Marschall, 1989) and that the p-center occurs earlier relative to the vowel as onset complexity increases (Ryan, 2014). This study provides the necessary acoustic evidence to support the claim that the timing of the intensity maximum covaries with the p-center across syllables of different onset complexities, and furthermore, it shows that the timing of the f<sub>0</sub> maximum is impacted in a qualitatively similar way. These results suggest that the timing of the intensity and f<sub>0</sub> maxima are acoustic factors that may influence gradient weight in English.

In stimuli controlled for stress and categorical weight, these results reflect general acoustic properties of onset complexity in English. Since the phonetic properties associated with onset complexity are not qualitatively similar to the primary phonetic properties associated with categorical weight (integrated intensity & duration), it suggests that the gradient weight afforded English syllables due to their onset characteristics is motivated semi-independently of acoustic cues to categorical weight. The semi-independence of gradient and categorical weight criteria are well captured by the interaction between categorical and scalar constraints in Ryan's (2011) Maximum Entropy (MaxEnt) grammar for gradient weight in quantitative verse. In MaxEnt grammar (Goldwater & Johnson, 2003), a type of Harmonic Grammar (Smolensky & Legendre, 2006), candidate harmonies are interpreted probabilistically, such that candidates with higher harmony are more likely outputs than those with lower harmonies; and at a given locus, categorical constraints are either violated or not, while scalar (gradient) constraints may be violated to some real-valued degree motivated by phonetics. Ryan (2011) used this framework to model the typology of gradient weight behavior he observed in the quantitative verse of several different languages, using categorical constraints to capture categorical weight behavior, and gradient constraints to capture gradient weight behavior. Although Ryan (2011) used log mean duration of the rime to motivate gradient constraints, the results of studies like this one could be used to motivate gradient constraints in future work.

Although Gordon's (2002) account does not predict the gradient weight of onsets in English, its notion of phonetic effectiveness 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 acoustic cues motivating gradient onset weight in English are typically subservient to the cues that motivate categorical weight. Only within a weight category, where cues to categorical weight are less effective to perceptually distinguish syllable structures, do cues to gradient weight exert a noticeable effect. Crucially, the acoustic motivators of gradient weight are those that are not sufficient to motivate categorical weight elsewhere in the language. For example, while rime duration and intensity alone are incapable of accounting for the gradient weight of onset complexity, they are perfectly capable of motivating categorical weight behavior elsewhere in the language (Delattre, 1966; Fry, 1955; Kochanski, Grabe, Coleman, & Rosner, 2005). By this logic, for the subset of the English lexicon with categorical O(R)V > RV weight, because the weight distinction between obstruent and sonorant onsets is categorical, we should expect to see duration and integrated intensity behave as predicted by Gordon (2002), such that obstruent onsets have greater rime duration and integrated intensity than sonorant onsets.

Although rime duration and integrated intensity play a role in the motivation of categorical weight in English, the effect of onset complexity on these measures is inconsistent with the fact that more onsets are probabilistically heavier than few onsets in the language. However, all other measures are still qualitatively consistent with the predictions of Gordon's (2002) account. In particular, for the difference in average rime intensity, the timing of the maximum intensity and  $f_0$ , and the timing of the maximum change in intensity and  $f_0$ , the difference between the zero- and one-onset words (1) is in the predicted direction for perceptual accounts of weight and (2) is greater than the difference between one- and multi-onset words. This second point echoes the crosslinguistic observation that CV > V weight is crosslinguistically extremely rare yet far more common than CCV > CV weight criterion, which are virtually unattested in categorical onset weight (Gordon, 2005). In this way, the acoustic motivators of gradient weight appear to contribute to weight in a manner consistent with the observed crosslinguistic typology of weight, but their contribution is not capable of carving up the perceptual space to the same extent as motivators of categorical weight.

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 in the narrower maximum width and longer, thicker tails of the violin plots associated with onsetless syllables compared to onsetful syllables in Figures 5, 7, and 8 (right). For example, the broader distributions of the timing of the intensity and  $f_0$  maxima 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 had no impact on the maximum steepness change 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 prevent the realization of a more gradual change in intensity for syllables with more complex onsets if it were accompanied by 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 intensity. The p-center tends to occur prior to the beginning of the vowel and the intensity maximum. The location of maximal change in intensity is a possible point which meets these temporal characteristics and involves calculation over the intensity domain. If it is the case that the p-center acts as the point after which intensity 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.

## 6. Conclusion

The current paper presents results from a production study of English monosyllabic words designed to show the impact of onset complexity on acoustic characteristics of the syllable and finds that onset weight effects in English are most likely attributable to acoustic correlates of the p-center, like the timing of the intensity peak, rather than to correlates of categorical weight, like integrated intensity. Given the difference of these results compared to those found in languages with categorical onset weight, these findings suggest that phonetic effects of onset complexity in English are independent of the phonological weight system but are exploited by it to enhance syllable prominence.

## 7. Acknowledgements

I would like to thank Yolanda Chow, Iliana De Dios, Natalia Reed, Lin Tian, and Joyee Wong for their generous assistance with data segmentation. This work has greatly benefited from discussion with Marc Garellek, Eric Baković, and Gabriela Caballero; the careful comments of two anonymous reviewers and the Associate Editor; and helpful feedback received at the Winter 2017 Meeting of the Acoustical Society of America where an earlier version of this project was presented. Any errors are my own.

## Notes

<sup>1</sup>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.

<sup>2</sup>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
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. doi:<https://doi.org/10.18637/jss.v067.i01>
- 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>
- Davidson, L., & Erker, D. (2014). Hiatus resolution in american english: The case against glide insertion. *Language*, 90(2), 482–514. doi:<https://doi.org/10.1353/lan.2014.0028>
- Davis, S. (1988). Syllable onsets as a factor in stress rules. *Phonology*, 5(1), 1–19. doi:<https://doi.org/10.1017/S0952675700002177>
- de Jong, K. J. (1994). The correlation of p-center adjustments with articulatory and acoustic events. *Perception & Psychophysics*, 56(4), 447–460. doi:<https://doi.org/10.3758/BF03206736>
- 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>
- Delattre, P. (1966). A comparison of syllable length conditioning among languages. *IRAL-International Review of Applied Linguistics in Language Teaching*, 4(1-4), 183–198. doi:<https://doi.org/10.1515/iral.1966.4.1-4.183>
- 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.
- 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.
- Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America*, 27(4), 765–768. doi:<https://doi.org/10.1121/1.1908022>
- Fry, D. B. (1958). Experiments in the perception of stress. *Language and speech*, 1(2), 126–152. doi:<https://doi.org/10.1177/002383095800100207>
- Fry, D. B. (1965). The dependence of stress judgments on vowel formant structure. In *Phonetic sciences* (pp. 306–311). doi:<https://doi.org/10.1159/000426965>
- Garellek, M. (2013). *Production and perception of glottal stops* (Doctoral dissertation, UCLA).
- Goedemans, R. (1998). Weightless segments. *The Hague: Holland Academic Graphics*.
- Goldwater, S., & Johnson, M. (2003). Learning OT constraint rankings using a maximum entropy model. In *Proceedings of the Stockholm workshop on variation within Optimality Theory*.
- 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(2), 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.
- Janker, P., & Pompino-Marschall, B. (1991). Is the p-center position influenced by ‘tone’? In *Proceedings of the 12th international congress of phonetic sciences*.
- Jones, E., Oliphant, T., Peterson, P., et al. (2001). SciPy: Open source scientific tools for Python. [Online; accessed <today >]. Retrieved from <http://www.scipy.org/>
- 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.
- Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *The Journal of the Acoustical Society of America*, 118(2), 1038–1054. doi:<https://doi.org/10.1121/1.1923349>
- 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.
- Moore, B. C. (2012). *An introduction to the psychology of hearing*. Brill.
- 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. [B.]. (1989). On the psychoacoustic nature of the p-center phenomenon. *Journal of phonetics*.
- R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. Retrieved from <https://www.R-project.org/>
- 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>
- Seyfarth, S., & Garellek, M. (2015). Coda glottalization in american english. In *Proceedings of the International Congress of Phonetic Sciences*.
- Smolensky, P., & Legendre, G. (2006). *The harmonic mind: From neural computation to optimality-theoretic grammar (cognitive architecture)*, vol. 1. MIT press.
- 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.

- Van Heuven, V., & Sluijter, A. M. (1996). Notes on the phonetics of word prosody. *Stress patterns of the world: Part 1*, 233–269.
- Villing, R. (2010). *Hearing the moment: Measures and models of the perceptual centre* (Doctoral dissertation, National University of Ireland Maynooth).
- Warren, R. M. (1970). Elimination of biases in loudness judgments for tones. *The Journal of the Acoustical Society of America*, 48(6B), 1397–1403. doi:<https://doi.org/10.1121/1.1912298>
- 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 1** Target Items. Listed by number of onset segments.

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

**Table 2** 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		