

What is the domain for weight computation: the syllable or the interval?*

Aron Hirsch

Massachusetts Institute of Technology

1 Introduction

It is well known that the distribution of lexical stress is sensitive to the weight of rhythmic units such that heavier units more strongly attract stress. The aim in this paper is to arbitrate between two approaches to defining the rhythmic unit relevant for weight computation. The traditional approach links weight to the **syllable**, with its internal constituency of onset and rime. Weight is computed over the **rime**; onset consonants either do not contribute weight (e.g. Halle & Vergnaud 1980, and later work reviewed in Blevins 1995), or make a gradient weight contribution subordinate to that of the rime (Kelly 2004, Gordon 2005, Ryan 2013). The alternative, recently proposed in Steriade (2012), is a non-syllable-based approach which takes the rhythmic unit to be the total vowel-to-vowel **interval**. Intervals do not have internal constituency, so the entire interval is the weight domain.

The paper is organized as follows. In section 2, I review the interval theory, and show that it differs from the syllable theory in the way it parses certain inter-vocalic consonants. In section 3, I report results of a production experiment designed to determine which parse is correct. Experimental results oppose the syllable theory, and are generally supportive of the interval theory. In section 4, alternative versions of the syllable theory that could account for some of the experimental data are considered, and evidence against them is provided. Section 5 reconciles the interval theory with quantitative results in the extant literature.

2 Intervals, and where they dissociate from syllables

2.1 What are intervals? The interval is defined as the string from the beginning of one nucleus to the beginning of the next nucleus, or the end of the domain. All post-nuclear consonants are parsed into the interval with the nucleus that precedes them. Consonants preceding the initial nucleus in the domain are not parsed into an interval. The string of consonants and vowels in (1) is parsed into intervals as shown (*' notates interval divisions):

(1) <C₁> V₁C₂* V₂C₃C₄* V₃C₅C₆C₇* V₄

The first interval begins with V₁ and goes to the beginning of V₂, so it includes V₁ and C₂; the second begins with V₂ and goes to the beginning of V₃, so includes V₂, C₃, and C₄; and so forth, ending with the fourth interval, which goes from the beginning of V₄ to the end of the domain. C₁, which precedes V₁, is not parsed into an interval.

2.2 How is weight computed in interval theory? I adopt from Gordon (e.g. 2002) the idea that weight is a scalar dimension, and, as a first approximation, link weight to acoustic duration (though other factors, like sonority, are also relevant). Intervals do not have internal constituency, so all segments within the interval

* I am grateful to Edward Flemming and Donca Steriade for much helpful advice and discussion. Thanks also to Michael Kenstowicz, Adam Albright, Michael Wagner, Juliet Stanton, and Paul Marty, as well as to the audiences at Phonology 2013 (UMass Amherst), RUMMIT 2013 (UMass Amherst), and the Spring 2013 Advanced Phonology class at MIT. The experiment was carried out in the prosody.lab at McGill University; thanks to Michael Wagner for making this possible, and to Elise McClay, Thea Knowles, Symon Jory, and Lauren Garfinkle for running participants. Thanks to Erin Olson and Juliet Stanton for help with data annotation. The author is supported in part by a doctoral fellowship from the Social Sciences and Humanities Research Council of Canada. All inadequacies are my sole responsibility.

contribute weight. The weight contribution of a given segment is commensurate with its duration. The weight of the interval is a function of the total duration of the interval: the greater the total duration, the heavier the interval. Gradient weight hierarchies like (2) are predicted:

$$(2) \quad \iota_1 V_1 C_1 C_2 C_3 > \iota_2 V_1 C_1 C_2 > \iota_3 V_1 C_1 > \iota_4 V_1$$

Each interval is a proper superset of the intervals to its left: ι_1 contains one more consonant than ι_2 , which contains one more consonant than ι_3 , which contains one more consonant than ι_4 , which contains just V_1 . I will make the assumption that the total duration of an interval positively correlates with the number of consonants it contains: the more consonants, the longer the interval. ι_1 is longer – and thus heavier – than ι_2 , which is longer and heavier than ι_3 , which is longer and heavier than ι_4 .

2.3 Dissociating syllables and intervals Syllables and intervals differently parse certain inter-vocalic consonants. Syllable parsing is onset maximizing, subject to constraints on complex onsets: inter-vocalic consonants are parsed in the onset of the syllable headed by the subsequent vowel, rather than the coda of the syllable headed by the preceding vowel. VCV syllabifies V.CV, never VC.V. VCCV syllabifies V.CCV if CC is a legal onset cluster, and otherwise VC.CV; the division is never VCC.V. As seen in (1), interval parsing goes the opposite way: inter-vocalic consonants are parsed in the interval of the preceding vowel, not the interval of the subsequent vowel. VCV is divided VC*V, and VCCV is divided VCC*V. The parses which never occur according to the syllable theory are the parses the interval theory predicts.

Different parsing on the two theories leads to different predictions about *whether* and if so, *where*, C in VCV and CC in VCCV contribute weight. Start with the syllable parse: V.CV and V.CCV (assuming legal onset CC). If weight is computed solely over the rime, C and CC in the onset make **no** weight contribution; if weight computation is onset-sensitive, C and CC contribute weight to the **final** syllable. On an interval parse, VC*V and VCC*V, C and CC make a weight contribution to the **initial** interval.

It is possible to use stress to diagnose how C and CC contribute weight. Stress is more strongly attracted to a vowel the heavier the rhythmic unit that vowel occurs in. We compare how strongly stress is attracted to the initial or final vowel in VCV vs. VCCV. If C and CC make no weight contribution, as strongly as stress is attracted initially or finally in VCV, it should be equally strongly attracted initially or finally in minimal VCCV. If C and CC contribute weight to the rhythmic constituent containing the final vowel, assuming that CC makes a greater weight contribution than C, stress should be attracted to the final vowel more strongly in VCCV than in VCV. If C and CC contribute weight to the rhythmic constituent containing the initial vowel, stress should be attracted more strongly to the initial vowel in VCCV than in VCV.

To see this concretely, compare stress predictions given syllable parsing and interval parsing for the bi-vocalic nonce pair *aka*, *akra*. First, the syllable parse:

$$(3) \quad \begin{array}{ll} \text{a. } a.k\bar{a} & V.CV \\ \text{b. } a.kra & V.CCV \end{array}$$

The initial syllable is identical in (a) and (b), being [a] in both cases. The final syllable is CV [ka] in (b) and CCV [kra] in (a). If C and CC do not contribute weight (rime-based weight computation), CV and CCV are of equivalent weight; the rime in both cases is V [a]. With an equivalent initial syllable and equivalent final syllable, stress will not be attracted more strongly initially or finally in (a) than in (b). If C and CC *do* contribute weight (onset-sensitive weight computation), CCV in (b) is heavier than CV in (a). With an equivalent initial syllable and heavier final syllable, stress will be attracted finally more strongly in (b).

Now, the interval parse:

$$(4) \quad \begin{array}{ll} \text{a. } ak^*a & VC^*V \\ \text{b. } akr^*a & VCC^*V \end{array}$$

The initial interval is VC [ak] in (a) and VCC [akr] in (b). As per the weight hierarchy in (2), VCC is assumed to be of greater duration and thus heavier than VC. The final interval is identical in (a) and (b), being V [a] in both cases. With a heavier initial interval and equivalent final interval, stress will be attracted to the initial vowel more strongly in (b).

To arbitrate between syllable theory and interval theory, this study will use a nonce word production experiment to determine how strongly stress is attracted initially or finally in VCV vs. VCCV sequences. Are participants as likely to produce initial vs. final stress with VCV as with minimal VCCV (syllables, rime-based)? Are they more likely to produce final stress with VCCV (syllables, onset-sensitive)? Or, are they more likely to produce initial stress with VCCV (intervals)?

2.4 Some initial evidence for intervals Before proceeding to the experiment, I discuss a couple of examples of experimental phonetic data already in the literature which support interval parsing over syllable parsing. The goal is to lend enough plausibility to interval theory to justify its consideration in the current study. Farnetani & Kori (1986) and McCrary (2004) report that the duration of a vowel is in a trading relationship with the duration of the entire consonantal interlude intervening between that vowel and the next vowel: V_1 is longer in V_1CV than in V_1CCV . On a syllable parse, this is surprising. $V_1.CV$ and $V_1.CCV$ or $V_1C.CV$ are the syllable divisions, meaning consonants not in the syllable with V_1 are having an effect on its duration. On an interval parse, the divisions are V_1C^*V and V_1CC^*V . The consonants affecting V_1 's duration are in the interval with V_1 and in so far as intervals tend towards uniform duration, the trading relationship is unsurprising: the greater the duration of the consonantal portion of the interval, the lesser the duration of the vowel to compensate. Note that McCrary's data show that compensation is never complete: CC adds more duration relative to C than shortening of the vowel compensates for. The weight hierarchy in (2) thus holds even when the trading relationship is taken into account. Other experimental data supportive of intervals show that perceived speech rate is affected by manipulation of interval duration, but not by manipulation of syllable duration (Kato et al. 2003). Additional arguments for intervals can be found in Steriade (2012).

3 Production experiment

3.1 Overview The goal of the experiment is to test dissociative predictions of intervals and syllables on the location of stress in VCV vs. VCCV sequences. Participants produce bi-vocalic nonce words. The consonantal interlude separating the initial and final vowels in the nonce words is varied between a single C, a CC cluster legal as a complex onset, and a CC cluster illegal as a complex onset. Legal clusters will be referred to as 'CC1' and illegal clusters as 'CC2'. The issues involved in defining legal onsets are discussed below in section 3.2, and in detail in section 4. Participants' productions are perceptually coded for whether they produced the nonce word with stress on the initial or final vowel.

The comparison between C and CC1 is the *aka* vs. *akra* comparison discussed in section 2.3. The syllable theory predicts no effect of C vs. CC1 on stress (rime-based), or a greater likelihood of final stress in CC1 than in C (onset-sensitive). The interval theory predicts a greater likelihood of initial stress in CC1.

Because the syllable theory potentially predicts a null effect for the C vs. CC1 comparison, the CC2 condition is added, introducing a new comparison for which the syllable theory does predict an effect: CC1 vs. CC2. The nonce word in CC2 is the same as in CC1 except that the order of the consonants in the medial cluster is permuted. This creates an illegal onset cluster. *akra* in CC1 transforms into *arka* in CC2; [rk] is not a legal onset. To see how the syllable theory predicts an effect for CC1 vs. CC2, consider the syllable parse of *akra* and *arka*:

- | | | | |
|-----|-----------------|-------|-----|
| (5) | a. <i>a.kra</i> | V.CCV | CC1 |
| | b. <i>ar.ka</i> | VC.CV | CC2 |

The initial syllable is V [a] in (a) and VC [ar] in (b). Coda consonants contribute weight, so VC is heavier than V. The final syllable is CCV [kra] in (a) and CV [ka] in (b). If weight is rime-based, CCV and CV are equivalent, both having the rime V; if weight is onset-sensitive, CV is lighter than CCV. With a heavier initial syllable in (b) and equivalent or lighter final syllable, stress will be attracted initially more strongly in (b).

The interval prediction for *akra* vs. *arka* is not entirely clear. Cluster phonotactics do not affect interval parsing, so the division is *akr*a* and *ark*a*, with [kr] and [rk] both in the initial interval. Even though they share a VCC profile and are made up of the same consonants, the initial intervals [kra] and [rka] are not, however, necessarily of the same weight. If the order of consonants in a cluster can affect the duration of the cluster, [kr] and [rk] may not be of equal duration; in that case, if differences in vowel duration do not

entirely compensate for differences in cluster duration, [akr] and [ark] would also not be of equal duration. Whichever is longer would be heavier, and stress would be attracted initially more strongly in that form. Durational measures are currently being extracted from the production data to see if there is a durational difference between [akr] and [ark]. The expectation is that [ark] will be of greater duration than [akr], in which case intervals would converge with syllables in predicting a greater occurrence of initial stress in *arka* than in *akra*. Pending the duration data, I leave the interval prediction tentative.

The syllable and interval predictions for C (e.g. k) vs. CC1 (kr) and CC1 vs. CC2 (rk) are summarized:

	Syllable theory	Interval theory
C vs. CC1	No effect, or final stress more likely in CC1	Initial stress more likely in CC1
CC1 vs. CC2	Initial stress more likely in CC2	

Syllable theory predicts (i) no effect of C vs. CC1, or a greater likelihood of final stress in CC1 than in C, coupled with (ii) a greater occurrence of initial stress in CC2 than in CC1. The clear prediction of interval theory is a greater occurrence of initial stress in CC1 than in C.

3.2 Stimuli Nonce words Items are triples of bi-vocalic nonce words differing minimally in the make-up of the medial consonantal interlude. The first condition has a single C, e.g. *keef*foos, ‘C condition’. In the second condition, a second C is added to create a legal onset cluster, e.g. *keef*loos, ‘CC1’. In the third condition, the order of the Cs in the cluster is permuted to give an illegal onset cluster, e.g. *keelf*foos, ‘CC2’.

For now, we will make the assumption that any cluster allowed word-initially is a legal onset and syllabified as such in medial position; this assumption will be re-visited in section 4. Items are grouped into three types. 6 items were constructed of each type.²

Group 1 (‘OL’): These items have either an obstruent other than [s] (3 items) or a liquid (3 items) in the C condition, an obstruent-liquid cluster in CC1, and a corresponding liquid-obstruent cluster in CC2. An example item is given in (6a), repeated from above. The full inventory of clusters from CC1 is given in (6b); all are allowed initially, and permute in CC2 to give clusters not allowed initially.

- (6) a. kee{f,l}oos C b. pl, tr, br, gr, fl, fr **OL items**
 keefloos CC1
 keelfoos CC2

Group 2 (‘sS’): These items have either [s] (3 items) or a sonorant (nasal, liquid) (3 items) in the C condition, an [s]-sonorant cluster in CC1, and a sonorant-[s] cluster in CC2. An example is in (7a), and the inventory of clusters from CC1 in (7b); each of the three clusters in (7b) occurred in two items.

- (7) a. kee{s,l}oo C b. sl, sm, sn **sS items**
 keesloo CC1
 keelsoo CC2

Group 3 (‘sT’): These items have either [s] (3 items) or a stop (3 items) in the C condition, an [s]-stop cluster in CC1, and a stop-[s] cluster in CC2. An example is in (8a), and the inventory of clusters from CC1 in (8b); each of the three clusters in (8b) occurred in two items.

- (8) a. wee{s,t}oof C b. sp, st, sk **sT items**
 weestoof CC1
 weetsoof CC2

² A fourth item type was also included in the experiment (for a total of 24 items), but will not be discussed here. The purpose of those items was to help understand results of an earlier experiment. They were like sT items (see below), but with an additional sonorant at the beginning of the consonantal interlude in all conditions. An example: *nel*{s,t}oof (C), *nel*tsoof (CC1), *nel*stoof (CC2).

Two notes regarding vowel quality. First: in all items, the initial vowel is tense. After lax vowels, single Cs are reported to be ambisyllabic or codas (e.g. Kahn 1976); I sought to avoid this effect. Second: the initial and final vowels are never reduced. When the nonce word is produced, one vowel receives primary stress and the other secondary.

Carrier sentence Nonce words are presented to participants orthographically as part of a carrier sentence:

- (9) I want to **keef**os.
 I want to **keefloos**.
 I want to **keelfoos**.

The carrier sentence is kept constant across conditions within an item. The carrier sentence did vary *between* items, though certain properties were invariant: in all carrier sentences, the nonce word was introduced as the last word in the sentence by infinitival *to*. This is important in two respects. First: the nonce word, occurring between unstressed *to* and the utterance boundary, can be produced with initial or final stress without a clash with an adjacent stress. Second: infinitival *to* identifies the lexical class of the nonce word as a verb. Across the English lexicon, there is a strong bias for bi-vocalic nouns to be stressed initially (92% of nouns in the Brown corpus are stressed initially, Sereno 1986), and bi-vocalic verbs to be stressed finally (85%). Pilot testing showed that the initial stress bias with nouns has a stronger effect on stressing of nonce forms than the final stress bias with verbs. A weaker bias allows effects of the phonological shape of the nonce word to emerge more easily with verbs.

Fillers In addition to the 24 experimental items, 24 fillers were created. These resembled experimental items, but the phonological profile of the nonce words was more widely varied in the fillers. Fillers came in four types: 6 fillers had a tri-vocalic nonce word (10a), 6 had a mono-vocalic nonce word (b), 6 had a bi-vocalic nonce word with a reduced penult (c), and 6 had a bi-vocalic nonce word with a reduced ultima (d).

- (10) a. Wow! I caught a **maykazoo**.
 b. Every Canadian has a **sarm**.
 c. I used a wrench to **viroor**.
 d. My friend invited me to **yoola**.

Filler trials were interspersed with experimental trials, and their goal was to prevent participants from falling into a repetitive prosody.

The full set of stimuli is available on my website: <http://prosodylab.org/lab/aron-hirsch/>.

3.3 Task Each trial involved three tasks, two of which we will be concerned with here: a production task, and a syllabification judgment task. Participants were presented with a carrier sentence containing a nonce word. They first did the production task: they were instructed to read the sentence silently to themselves until they felt comfortable to say it aloud, and then to record themselves saying the sentence aloud. Then came the syllabification judgment task: participants were asked to choose between two or three alternative syllabifications of the nonce word, presented in multiple-choice format. The syllabification judgment task will be discussed in detail in section 4.³

The experiment was run in a Latin Square design. There were 24 experimental trials, one per item. A given participant saw one of the three conditions from a given item (either C, CC1, or CC2), and an equal number of each of the three conditions across trials. The 24 experimental trials were alternated with the 24 filler trials. The experiment was run in a sound attenuated booth, and recordings for the production task were made using a Logitech USB headset.

3.4 Participants 28 volunteers participated in the study, most of them undergraduate students. All participants identified themselves as native speakers of North American English. Participants were run in the prosody.lab at McGill University, and received monetary compensation (\$10/hour).

³ The third task asked participants to judge the morphological complexity of the nonce word: if the nonce word were a real word, do they think it would be a simple word or a compound? For reasons of space, I cannot discuss this here.

3.5 Data annotation Three annotators (the author, and two colleagues) listened to all recorded sound files from the production experiment, and for each, coded whether they heard primary stress in the production of the nonce word on the initial or final vowel. Reported data combine the three annotators' coding in the following way. On trials where annotators were in agreement (73% of non-excluded trials), that judgment is recorded. On trials where there was disagreement, the recorded judgment is determined by majority vote. If two annotators coded stress as initial and one as final, the recorded judgment is initial; if two annotators coded stress as final and one as initial, the recorded judgment is final. The resultant data set was analyzed in a logit mixed effects model with stress (initial, final) as the dependent measure, condition (C, CC1, CC2) and item type (OL, sS, sT) as fixed effects, and random effects for item and participant that included random slopes for the fixed effects.

3.6 Excluded data 21% of sound files were excluded from analysis, according to the following criteria:

1. Sound files were excluded if the recording equipment failed, or if the participant's production was disfluent due to a false start, hesitation, coughing, laughing, or production of some other non-speech sound.
3. Annotators had the option of coding stress in the nonce word as "unclear" if they had difficulty discerning whether stress was initial or final. Sound files for which any of the three annotators coded "unclear" were excluded.
4. Because nonce words were presented to participants orthographically, there was variability in how participants pronounced the nonce words. Data were excluded if the participant gave an unexpected rendering of the medial cluster, e.g. [fuznaiv] for [fusnaiv], pronounced a vowel as lax when it was supposed to be tense, e.g. [witsuf] for [witsuf], or pronounced a vowel as reduced, e.g. [nuplar] for [nuplar]. These are the off-target pronunciations most likely to interact with stress placement. Other off-target pronunciations were included for practical reasons, due to their frequency. The most common was pronunciation of a target tense vowel as tense but with off-target quality, e.g. [fusniv] for [fusnaiv]. This does not affect the phonological template of the nonce word, [fusniv] = [fusnaiv] = CVVCCVVC, so should have minimal impact on the validity of inter-condition comparisons.

3.7 Overall results Figure 1 shows the relative proportion of trials in which participants produced the nonce word with initial vs. final stress in each of the three conditions, C, CC1, and CC2. Data from all participants and all item types – OL, sS, and sT – are included in the plot. Results show significant effects for both the C vs. CC1 comparison and the CC1 vs. CC2 comparison. Participants are more likely to produce the nonce word with initial stress in CC1 than C ($|z| = 2.36$), and in CC2 than CC1 ($|z| = 2.03$).

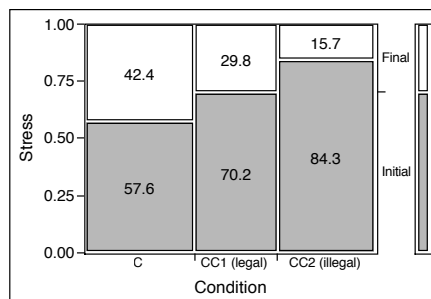


Figure 1: The relative proportion of initial vs. final stress by condition, all data.

3.8 Discussion The C vs. CC1 effect is inconsistent with syllables and supportive of intervals. Syllables predict no effect of C vs. CC1, or a greater occurrence of final stress in CC1 than in C. The actual result is a greater occurrence of initial stress in CC1 than in C — which is the interval prediction. C and CC are parsed in the initial interval; CC contributes more weight than C, so pulls stress more strongly to the initial vowel. The result for CC1 vs. CC2 – a greater occurrence of initial stress in CC2 – is consistent with syllables, but

is difficult to interpret as a syllable effect given that the C vs. CC1 effect contradicts syllables. Intervals can potentially unify the C vs. CC1 and CC1 vs. CC2 results. If, as speculated in section 2.3, the CC1 vs. CC2 manipulation interacts with interval duration such that the initial interval in CC2 has greater duration and is thus heavier, it follows from intervals that stress would be pulled to the initial vowel more strongly in CC2. Syllable theory cannot account for all of the data; depending on duration in CC1 and CC2, interval theory can.⁴

4 Syllabification

4.1 The problem Referring back to sections 2.3 and 3.1, the syllable predictions generated for the C vs. CC1 and CC1 vs. CC2 comparisons are based on the assumption that the syllable parse is V.CV in the C condition, V.CCV in CC1, and VC.CV in CC2. The V.CV division in the C condition and the VC.CV division in CC2 are uncontroversial under the syllable theory — but it is less clear whether all clusters in CC1 are syllabified V.CCV.

Recall that three types of clusters were classified as CC1: clusters with an obstruent other than [s] followed by a liquid in OL items, clusters with an [s] followed by a liquid or nasal in sS items, and clusters with an [s] followed by a stop in sT items. All of these clusters are allowed word-initially, but it may be, contra our earlier assumption, that a cluster being allowed word-initially is not a perfect predictor of its medial syllabification (Lowenstamm 1981, Steriade 1982, Clements & Keyser 1983). Some clusters allowed word-initially may be syllabified medially as a coda-onset sequence.

If CC1 included a mixture of V.CCV and VC.CV parsing, the syllable theory could account for the overall data in Figure 1. I spell out how in sections 4.1.1 and 4.1.2. To fairly test the syllable theory according to the above predictions, we need to consider a subset of the data in which all clusters included in CC1 are for sure tauto-syllabic.

4.1.1 Predictions with VC.CV in CC1 To see how the syllable theory could account for experimental results, we first need to generate the syllable predictions for C vs. CC1 vs. CC2 with hetero-syllabic parsing in CC1. Consider the syllable parse of *asa* (C) vs. *asta* (CC1) vs. *atsa* (CC2), assuming [st] in CC1 to be hetero-syllabic. First, C vs. CC1:

- (11) a. *a.sa* V.CV C
 b. *as.ta* VC.CV CC1

[s] is parsed in the coda of the initial syllable in CC1, which results in a heavier initial syllable in CC1 than in C, since coda consonants contribute weight, and [s] is parsed in the final onset in the C condition. VC [as] in (b) is a heavier initial syllable than V [a] in (a). The final syllable is CV in both (a), [sa], and (b), [ta]. With a heavier initial syllable in (b) and (near) equivalent final syllable, stress will be attracted to the initial vowel more strongly in (b). This is convergent with the interval prediction.

Now, CC1 vs. CC2:

- (12) a. *as.ta* VC.CV CC1
 b. *at.sa* VC.CV CC2

With hetero-syllabic parsing in CC1, parsing is hetero-syllabic in both CC1 and CC2. The initial syllable is VC in (a), [as], and (b), [at]. The final syllable is CV in both (a), [ta], and (b), [sa]. With a (near) equivalent initial syllable and (near) equivalent final syllable in (a) and (b), stress will be attracted no more strongly initially or finally in one than in the other. No effect of CC1 vs. CC2 is predicted.

4.1.2 How syllable theory could explain the overall data The original predictions with tauto-

⁴ Convergent data in Olejarczuk & Kapatinski (2014): participants are more likely to stress a tri-vocalic nonce word on the penult rather than the antepenult when the penultimate vowel is followed by a CC1 cluster than a single C, and a CC2 cluster than a CC1 cluster. The interval interpretation: the penultimate interval is heavier in CC1 (*tam*apr*ish*) than C (*tam*ap*ish*), and if CC2 clusters are longer than CC1, in CC2 (*tam*amp*ish*) than CC1 (*tam*apr*ish*).

syllabic parsing in CC1 are (i) no effect of C vs. CC1, or a greater occurrence of final stress in CC1 than in C, and (ii) a greater occurrence of initial stress in CC2 than in CC1. The predictions with hetero-syllabic parsing in CC1 are (i) a greater occurrence of initial stress in CC1 than in C, and (ii) no effect of CC1 vs. CC2. The experimental results bear out the tauto-syllabic prediction for C vs. CC1 (more initial in CC1), and the hetero-syllabic prediction for CC1 vs. CC2 (more initial in CC2). Suppose that parsing is by syllable, and that some clusters in CC1 are parsed as tauto-syllabic and some as hetero-syllabic. The hetero-syllabic clusters could be driving the C vs. CC1 effect, and the tauto-syllabic clusters could be driving the CC1 vs. CC2 effect. Results would then be explained in a way consistent with the rhythmic unit being the syllable, rather than the interval.

In the remainder of this section, I present two versions of the syllable theory which do predict that some clusters in CC1 are tauto-syllabic and some are hetero-syllabic, and thus could account for the overall results. I give evidence against both versions by considering subsets of the data. Each version makes a specific prediction about *which* clusters in CC1 are tauto-syllabic, and *which* are hetero-syllabic. To test a given version, I consider the subset of data where that version predicts tauto-syllabic parsing in CC1, and see whether the original syllable predictions hold in that subset. With only tauto-syllabic clusters in CC1, they should – but in neither case, does it seem that they do.

4.2 Alternative version of syllable theory: sonority-based First consider a version of the syllable theory in which the way a cluster is syllabified medially is entirely a function of its sonority profile. Clusters are syllabified medially as complex onsets only if they have a rising sonority profile; clusters with a flat or falling sonority profile are syllabified as a coda-onset sequence. This is regardless of whether the cluster is or is not allowed word-initially. Sonority-based syllabification predicts a divergence between medial syllabification of OL and sS clusters on the one hand, and sT clusters on the other hand. OL and sS clusters have a rising sonority profile, so are predicted to be tauto-syllabic, V.OLV and V.sSV; sT clusters have a flat or slightly falling sonority profile, so are predicted to be hetero-syllabic, Vs.TV.

I exclude sT items from analysis. Without sT items, the only clusters in CC1 are OL and sS, which are predicted to be tauto-syllabic, so the original syllable predictions should hold without complication: no effect of C vs. CC1 or a greater occurrence of final stress in CC1 than in C, coupled with a greater occurrence of initial stress in CC2 than in CC1. Does the C vs. CC1 effect observed in the overall data disappear or reverse with sT items excluded?

Figure 2 shows stress by condition including only trials with OL and sS items; sT items are excluded from all conditions.

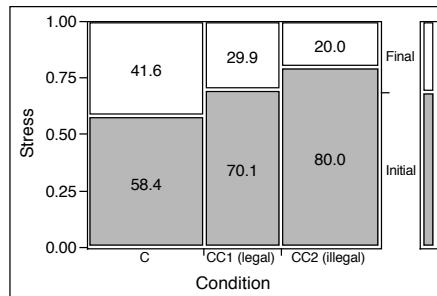


Figure 2: The relative proportion of initial vs. final stress by condition, OL and sS items only; sT items excluded.

Considering subsets of the data reduces statistical power, so no effects reach significance, but the trends observed in Figure 1 for C vs. CC1 and CC1 vs. CC2 *both* replicate in Figure 2. Participants are more likely to produce the nonce word with initial stress in CC1 than in C ($|z| = 1.53$), and in CC2 than in CC1 ($|z| = 1.33$). As a greater occurrence of initial stress in CC1 than in C is contrary to the original syllable predictions, Figure 2 provides suggestive evidence against this version of the syllable theory.

4.3 Alternative version of syllable theory: intuition-based The previous versions of the syllable theory we have considered categorically predict for a given cluster whether it is syllabified as a complex onset or a coda-onset sequence. This section presents results from the syllabification judgment task run concurrently

with the production task; these data show that speakers' intuitions about syllabification are variable, rather than categorical (see also Treiman et al. 1992). I consider a version of the syllable theory which grants that phonologically relevant syllabification is accessible to introspective judgment, and thus that variability in syllabification intuitions reflects variability in phonologically relevant syllabification. The phonological grammar makes available both V.CCV and VC.CV representations with the choice between them influenced by various factors (sonority profile, among others). The result is that syllabification is variable by participant and even by trial. To ensure tauto-syllabic parsing in CC1, I look at participants' syllabification judgments trial by trial and exclude those CC1 trials where the participant reports a hetero-syllabic judgment. I test whether the original syllable predictions hold in the resulting subset of data.

4.3.1 The syllabification judgment task As part of each trial, after the production task, participants were asked to judge the syllabification of the nonce word. They chose between two or three alternative syllabifications in a forced choice paradigm. The instruction was to select "the division of the fake word into parts which seems most natural based on how the word sounds". In the C condition, two options were presented, one with C syllabified in the onset of the final syllable (13a), and one with C syllabified in the coda of the initial syllable (13b). In CC1 and CC2, three options were presented, one with CC parsed as a complex onset (14a), one with CC as a coda-onset sequence (14b), and one with CC as a complex coda (14c).

- | | | | | | |
|------|---------------------------------|------|---|--|--|
| (13) | a. KEE/FOOS C
b. KEEF/OOS | (14) | a. KEE/FLOOS CC1
b. KEEF/LOOS
c. KEEFL/OOS | | KEE/LFOOS CC2
KEEL/FOOS
KEELF/OOS |
|------|---------------------------------|------|---|--|--|

The syllabified nonce words were written in capital letters in a way that indicated the sound of the word; for example, "KEE/PLAHR" was written rather than "kee/plar". This was intended to reinforce divisions chosen on the basis of sound (cf. Côté & Kharlamov 2011).

Syllabification in C and CC2 is uncontroversial: in C, an onset parse should be preferred over a coda parse; in CC2, a hetero-syllabic parse should be preferred over either tauto-syllabic parse. C and CC2 serve as control conditions to verify that the task is working properly. The case of interest was CC1.

4.3.2 Results As expected, participants virtually categorically intuited V.CV syllabification in the C condition, and VC.CV syllabification in CC2. The data are plotted in Figures 3 and 4 respectively, aggregated for all item types.

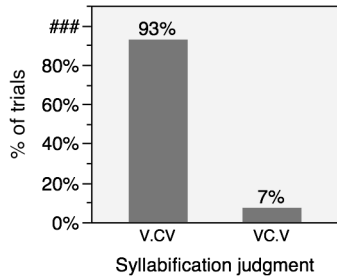


Figure 3: % of trials from the C condition in which participants' syllabification judgment was V.CV vs. VC.V.

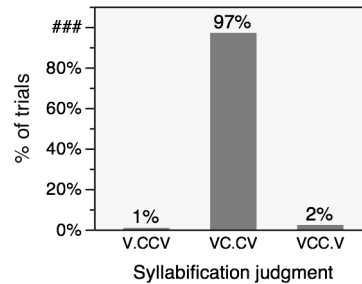


Figure 4: % of trials from the CC2 in which participants' syllabification judgment was V.CCV vs. VC.CV vs. VCC.V

Preferred syllabification in CC1 is presented separately for each item type in Figure 5 (next page); data from OL items are in (a), from sS items in (b)⁵, and from sT items in (c). With OL clusters, a complex onset parse is preferred: participants selected V.OLV as the most natural division in 85% of trials. Intuitions for sC clusters are highly variable. sS clusters are most commonly complex onsets, and sT clusters are most commonly coda-onset sequences, but these preferences are slight. The other parse is selected in both cases in >40% of trials.

⁵ No participant gave a VCC.V judgment in any trial for an sS item, so this category does not appear in the plot in (b).

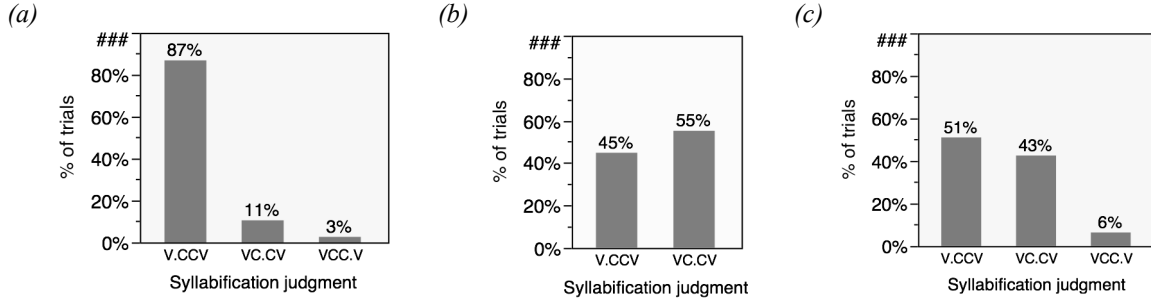


Figure 5: % of trials from CC1 where syllabification was *V.CCV* vs. *VC.CV* vs. *VCC.V* by item type; OL items (a), sS items (b), sT items (c).

4.3.3 Stress To avoid hetero-syllabic parsing in CC1, I exclude from analysis trials of all item types where the participant intuited a hetero-syllabic parse. With the few cases of coda parsing in the C condition and complex onset parsing in CC1 and CC2 also excluded, Figure 6 results. This shows stress by condition for C trials with intuited *V.CV* syllabification, CC1 trials with intuited *V.CCV* syllabification, and CC2 trials with intuited *VC.CV* syllabification.

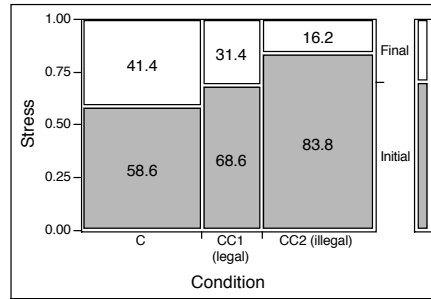


Figure 6: Stress by condition for C trials with a *V.CV* syllabification judgment, CC1 trials with a *V.CCV* syllabification judgment, and CC2 trials with a *VC.CV* syllabification judgment only.

With parsing in CC1 tauto-syllabic according to the current version of the syllable theory, the original syllable predictions should hold — and again are not fully borne out. The trends in Figures 1 and 2 replicate for both C vs. CC1 and CC1 vs. CC2. There is a nearly significant trend between C and CC1 such that participants are more likely to stress the initial vowel in CC1 ($|z| = 1.81$). For CC1 vs. CC2, there is a non-significant trend towards a greater occurrence of initial stress in CC2 ($|z| = 1.02$). The greater occurrence of initial stress in CC1 than in C is evidence against the syllable prediction.

4.4 OL items All versions of the syllable theory we have discussed converge in predicting one cluster type in CC1 to be tauto-syllabic: OL clusters. OL clusters are allowed word-initially (section 3.2), they have a rising sonority profile (section 4.2), and they are dominantly tauto-syllabic by intuition (section 4.3). The most uncontroversial test of the original syllable predictions would, therefore, be to exclude all sC items from analysis, and consider the subset of data including only trials with OL items.⁶ When this is done, no effects of condition are observed, not even as non-significant trends. This is not consistent with syllables or intervals: no effect of C vs. CC1 is consistent with syllables, but syllables predict more initial stress in CC2 than in CC1; intervals predict more initial stress in CC1 than in C. Both predict an effect *somewhere*.

Why might there be no differences between conditions in the OL items? Part of the reason might be noise in the items. It seems that, in general, some items bias more towards one stress pattern or the other in particular conditions for reasons independent of phonological template. Noise of this sort would be magnified

⁶ Recall from section 3.5 that the model used to analyze the overall data tested for an interaction with item type: results show a near significant interaction for OL items for the C vs. CC1 comparison, $|z| = 1.86$, and a significant interaction for the CC1 vs. CC2 comparison, $|z| = 1.97$.

when only one item type is considered because of the small number of items, and it could be that noise in the OL items happens to have the effect of cancelling inter-condition differences. At least for CC1 vs. CC2, though, there is likely something more systematic going on. One observation is that CC2 forms like *keelfoos* contain a sequence [ilf] that is phonotactically impossible in mono-morphemes. This could motivate participants to stress them as two words, with the nuclear stress rule then assigning final stress. This would cancel any effect of more initial stress in CC2 that could come from intended mono-morphemic parses. The CC2 forms in other item types, e.g. *veeksay*, may not require the two-word parse. Further experimental work is needed to properly understand the OL data.⁷

4.5 Summary I have provided evidence against three versions of the syllable theory. The first version (licit initially = complex onset medially, section 3.2) is ruled out by the overall data, Figure 1. The second version (rising sonority = complex onset medially, section 4.2) is contradicted by trends in the subset of data with OL and sS items, Figure 2. The intuitive version (section 4.3) is contradicted by trends in the subset of data with the intuited parse in CC1 tauto-syllabic, Figure 6. All plots have shown a trend towards a greater occurrence of initial stress in CC1 than in C — as the interval theory predicts.

5 Reconciling intervals with previous quantitative results

I now reconcile the results of the current experiment with those of previous experimental and corpus work comparing how strongly stress is attracted to the initial vs. final vowel in bi-vocalic forms when the number of **word-initial** consonants is manipulated; the present study manipulates consonants **word-medially**. I show how a modification of interval theory can account for all of the data in a unified way.

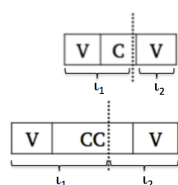
Word-initial consonants make a weight contribution to attract stress to the initial vowel. Kelly (2004) experimentally compared stressing of nonce bi-vocals like *bontoon* vs. *brontoon*, manipulating whether there was a single C [b] or CC cluster [br] word-initially. Participants were asked to judge whether, according to their intuition, the nonce word would be stressed initially or finally. A judgment of initial stress was more frequent in *brontoon* than *bontoon*. Kelly replicated the effect in corpus work. He showed that, across the English lexicon, there is a positive correlation between the number of initial consonants in a bi-vocalic word and the likelihood that the word is initially stressed: initial stress is more likely with CCC initially (*stratum*) than with CC initially (*blossom*) than with C initially (*canal*) than with no consonant initially (*abyss*). Ryan (2013) found the same effect experimentally, and in the Russian lexicon.

Kelly and Ryan interpret their data in syllable-based terms. Word-initial consonants are parsed in the onset of the initial syllable; they propose that onset consonants make a gradient weight contribution. Ryan captures this by having the left edge of the weight domain not be the beginning of the rime, but instead be the perceptual centre (p-centre) of the syllable. The p-centre is an acoustic even which, if the syllable were said over and over with a metronome, would align with the downbeat of the metronome. The p-centre of a CV syllable is roughly at the beginning of the rime, but the p-centre is displaced left from that point to an extent correlating with the duration of additional onset consonants. The p-centre in a CCV syllable is further left than in a CV syllable, and the p-centre in a CCCV syllable is further left still (diagrammed). So, the greater the number of word-initial consonants, the more consonants in the onset of the initial syllable, the heavier the initial syllable, and the more strongly stress is attracted to the initial vowel.

On the interval theory, consonants prior to the initial vowel are not parsed in an interval, so are not predicted to make a weight contribution. It is, however, possible to adapt intervals in much the same way Ryan did with syllables, with recourse to a p-centre-like effect. Suppose that the boundary of an interval is not fixed at the beginning of a vowel, but is displaced left from the beginning of the vowel as a function of the duration of preceding consonants. The more word-initial consonants, the further left the boundary of the initial interval, the more pre-vocalic consonantal duration included in the interval, the heavier the interval, and the more strongly stressed is attracted to the initial vowel.

Crucially, even when augmented with a p-centre-like effect, interval theory can account for the results of the current study — though the situation is a bit more complicated than before. Without p-center effects, C in

⁷ Pending further data, I cannot rule out a version of syllable theory where syllabification intuitions do not reveal phonologically relevant syllabification, and only OL clusters are in fact tauto-syllabic medially.



VCV and CC in VCCV were entirely parsed in the initial interval, VC*V and VCC*V. CC contributed more weight to the initial interval than C, and stress was attracted to the initial vowel more strongly with CC. With p-center effects, the boundary between the initial and final intervals is displaced left from the beginning of the final vowel as a function of the duration of consonants preceding the final vowel, so the medial consonantal interlude is partially parsed in the initial interval and partially parsed in the final interval (diagrammed). Assuming that consonants cannot shift left the interval boundary by more than half of their duration, however, medial consonants contribute *more* of their weight to the initial interval than to the final interval, and the net weight gain with CC over C is thus to the initial interval. Stress is still predicted to be attracted to the *initial* vowel more strongly with CC.

It thus follows from the interval theory augmented with a p-centre-like effect that word-initial consonants (Kelly, Ryan) and word-medial consonants (current study) both pull stress to the initial vowel. Word-initial consonants are partially parsed in the initial interval and otherwise are not parsed in any interval, so they make their sole weight contribution to the initial interval. Medial consonants contribute some weight to the final interval, but are otherwise parsed in the initial interval, and make their greater weight contribution to the initial interval. Previous and current results are accounted for.

6 Conclusion and outlook

I have provided experimental evidence in opposition to the syllable as the rhythmic unit relevant for weight computation. Although syllables have been evoked for purposes other than weight computation, the present study should serve as a stimulus to re-evaluate our understanding of rhythmic constituency. Results are supportive of the recently proposed interval theory. In production of bi-vocalic nonce words, the likelihood of initial stress increases when there is a CC cluster medially than when there is just a single C; in the interval theory, medial consonants are parsed in the interval of the initial vowel, so contribute weight to pull stress to the initial vowel, CC contributing more weight than C. The next step is to directly correlate stress with interval duration. To do this, acoustic measures of duration must be extracted from the production data. If the interval is the right rhythmic unit, the degree to which stress is attracted initially should positively correlate with the duration of the initial interval across conditions and item types.

References

- Blevins, J. (1995). The syllable in phonological theory. In John Goldsmith (ed.) *Hndbk. Phon. Th.*, Blackwell, 204-244.
- Clements, G. & S. Keyser. (1983). *CV phonology: a generative theory of the syllable*. LI Mon. 9, Cambridge: MIT Press.
- Côté, M.-H. & V. Kharlamov. (2011). The impact of experimental task on syllabification judgments: a case study of Russian. In C. Cairns & E. Raimy (eds.) *Hndbk. of the Syllable*, Leiden: Brill, 273-294.
- Farnetani, E. & S. Kori. (1986). Effects of syllable and word structure on segmental durations in spoken Italian. *Speech Communication* 5(1), 17-34.
- Gordon, M. (2005). A perceptually-driven account of onset-sensitive stress. *Natural Lang. and Ling. Theory* 23, 595-653.
- Gordon, M. (2002). A phonetically-driven account of syllable weight. *Language* 78, 51-80.
- Halle, M. & J.-R. Vergnaud. (1980). Three dimensional phonology. *Journal of Linguistic Research* 1, 83-105.
- Kato, H, T. Minoru, & Y. Sagisaka. (2003). Functional differences between vowel onsets and offsets in temporal perception of speech: Local-change detection and speaking-rate discrimination. *J. Acoust. Soc. Am.* 113(6), 3379-89.
- Kelly, M. H. (2004). Word onset patterns and lexical stress in English. *Journal of Memory and Language* 50, 231-244.
- Khan, D. (1976). *Syllable-based generalizations in English phonology*. New York: Garland.
- Lowenstamm, J. (1981). On the maximal cluster approach to syllable structure. *Linguistic Inquiry* 12, 575-604.
- McCrary, K. (2004). *Reassessing the role of the syllable in Italian phonology: An experimental study of consonant cluster syllabification, definite article allomorphy, and segment duration*. Ph.D. dissertation, UCLA.
- Olejarczuk, P. & V. Kapatinski. (2014). The syllabification of medial clusters: evidence from stress assignment. Poster from LSA Annual Meeting, Minneapolis.
- Ryan, K. (2013). Onsets contribute to syllable weight: statistical evidence from stress and meter. Ms., Harvard.
- Sereno, J. (1986). Stress pattern differentiation of form class in English. *J. Acoust. Soc. Am.* 27: S36.
- Steriade, D. (2012). Intervals vs. syllables as units of linguistic rhythm. Handouts, EALING, Paris.
- Steriade, D. (1982). *Greek prosodies and the nature of syllabification*. Ph.D. dissertation, MIT.
- Treiman, R., J. Gross, & A. Cwikel-Glavin. (1992). The syllabification of /s/ clusters in English. *J. of Phon.* 20, 383-402.