ONSETS CONTRIBUTE TO SYLLABLE WEIGHT: STATISTICAL EVIDENCE FROM STRESS AND METER

Kevin M. Ryan

Harvard University

While some accounts of syllable weight deny a role for onsets, onset-sensitive weight criteria have received renewed attention in recent years (e.g. Gordon 2005, Topintzi 2010). This article presents new evidence supporting onsets as factors in weight. First, in complex stress systems such as those of English and Russian, onset length is a significant attractor of stress both in the lexicon and in nonce probes. This effect is highly systematic and unlikely, it is argued, to be driven by analogy alone. Second, in flexible quantitative meters (e.g. in Sanskrit), poets preferentially align longer onsets with heavier metrical positions, all else being equal. A theory of syllable weight is proposed in which the domain of weight begins not with the rime but with the p-center (perceptual center) of the syllable, which is perturbed by properties of the onset. While onset effects are apparently universal in gradient weight systems, they are weak enough to be usually eclipsed by the structure of the rime under categorization. This proposal therefore motivates both the existence of onset weight effects and the subordination of the onset to the rime with respect to weight.*

Keywords: stress, weight, syllable, onset, metrics, p-center

1. Introduction. A variety of phonological phenomena invoke syllable weight distinctions, the two most discussed being weight-sensitive stress and quantitative poetic meter. In the former, stress placement varies according to the distribution of syllable weight in the word; in the latter, the distribution of heavy and light syllables is regulated in verse constituents. Weight is often treated by the grammar as binary (heavy or light), though more articulated scales are also possible (see e.g. Hayes 1995, Morén 2001, Gordon 2002, 2006, de Lacy 2004, Ryan 2011a). As is well established, syllable weight criteria typically refer only to the structure of the rime (i.e. nucleus and coda), ignoring the onset. Thus, it has often been assumed that onsets are incapable of contributing to weight (e.g. Halle & Vergnaud 1980, Hyman 1985, Hayes 1989, Goedemans 1998, Morén 2001). Nevertheless, apparent cases of onset-sensitive criteria, though uncommon, have been accumulating since the 1980s. The following paragraph provides a brief synopsis of some such cases (following surveys in Davis 1988, Goedemans 1998, Hajek & Goedemans 2003, Gordon 2005, and Topintzi 2010).

Among stress systems, a V < CV criterion (i.e. a null onset is lighter than a filled one) is claimed for the Australian languages Agwamin, Alyawarra, Aranda, Kaytetj, Kuku-Thaypan, Lamalama, Linngithig, Mbabaram, Parimankutinma, Umbindhamu, Umbuygamu, and Uradhi, for the Amazonian languages Pirahã (Muran), Juma (Tupian), and Banawá (Arawan), and for Manam (Austronesian), Nankina (Papuan), and Iowa-Oto (Siouan) (see Goedemans 1998:249, Gordon 2005, and Topintzi 2010 for references; see also Gahl 1996, Downing 1998, and §5.3 below). V < CV is arguably also implied by stress-conditioned epenthesis in Ainu and Dutch (Booij 1995:65, Topintzi 2010:63). CV < C:V (i.e. a geminate onset is heavy) is claimed for at least Bellonese,

* I wish to thank Adam Albright, Angela Carpenter, Stanley Dubinsky, Rob Goedemans, Matthew Gordon, Maria Gouskova, Dieter Gunkel, Bruce Hayes, Marek Majer, Donca Steriade, Nina Topintzi, Kie Zuraw, two anonymous referees, and audiences at Harvard, MIT, Stanford, UCLA, UMass-Amherst, the University of Delaware, the 2013 LSA annual meeting, and WCCFL 28 for their questions, pointers, and criticisms regarding various aspects of this material. Discussions with Bruce Hayes were particularly formative as I set about exploring this topic. It goes without saying that I am alone responsible for the shortcomings of the article and any errors it might contain.

Marshallese, Pattani Malay, and Trukese (Topintzi 2010). Gordon (2005) cites two languages (Bislama and Nankina) as observing CV < CCV, though both are questionable (Topintzi 2010:223). Onset quality can also condition stress placement, as in Pirahã (Everett & Everett 1984, Everett 1988, Gordon 2005), Arabela (Payne & Rich 1988), Tümpisa Shoshone (Dayley 1989), Karo (Gabas 1999, Topintzi 2010:39; cf. Blumenfeld 2006), and Puluwat (Elbert 1972, Goedemans 1998:142), a frequent generalization being that voiceless obstruent onsets are heavier than other onsets. Beyond stress, onsets are also claimed to be invoked in cases of compensatory lengthening and prosodic minimality (Beltzung 2008, Topintzi 2010).

Almost all previous research on onset weight (Kelly 2004 being an exception; see §2.1) addresses the question from the perspective of categorical weight criteria, such as in the languages just enumerated. This article considers a complementary empirical field, namely, stress systems and meters exhibiting gradient variation. Analysis of these systems and of related experimental data permits syllable weight to be put under a more powerful microscope than the analysis of categorical criteria permits. For example, while binary weight in English is rime-based, onsets are significant as statistical predictors of stress placement in both existing (§2.1) and novel (§3.1) words. Similarly, while binary weight in Sanskrit ignores the onset, Sanskrit poets exhibit significant sensitivity to the onset in aligning syllables to metrical templates (§4.1–4.2). In every case, longer onsets pattern as heavier.

To account for these onset effects, it is proposed that the domain over which syllable weight is computed in stress and meter begins not with the rime but with the p-center (perceptual center), an event marking the perceived downbeat of the syllable (§5). While p-centers generally approximate the left edge of the rime, they are perturbed as a function of the onset, tending, for instance, to occur earlier in syllables with longer onsets. Longer onsets are therefore predicted to augment the percept of heaviness. Crucially, however, this approach also predicts onset segments to exert a weaker influence on weight than rime segments, given that the former are parsed only partially (if at all) into the weight domain, while the latter contribute in their entirety. Assuming with Gordon (2002) that weight categorization seeks to maximally discriminate syllables in a perceptual space, onset-based criteria are predicted to be typically inferior to rime-based criteria under (especially binary) categorization, explaining the observed onset-rime asymmetry.

The article is organized as follows. Stress is considered in §2, first in English and then in Russian (with a note on Italian). Section 3 treats experimental evidence for productivity and the role of analogy. Poetic meter is analyzed in §4, with a focus on three Vedic meters, Epic Sanskrit, and Kalevala Finnish. Lastly, §5 proposes an explanation and generative model and §6 concludes.

2. Onset effects in stress.

2.1. English. Generative treatments of English stress (e.g. Chomsky & Halle 1968, Halle & Keyser 1971, Halle 1973b, Liberman & Prince 1977, Hayes 1982, Halle & Vergnaud 1987, Kager 1989, Halle & Kenstowicz 1991, McCarthy & Prince 1993, Burzio 1994, Hammond 1999, Pater 2000) have generally assumed that onsets are inert as determinants of stress placement (though see Nanni 1977 on adjectives in *-ative*, in which stress is apparently sensitive to the obstruency of the immediately preceding onset). Kelly (2004), however, finds that the number of consonants in the initial onset of an English disyllable positively correlates with its propensity for initial primary stress. This correlation is significant not only in the aggregate, but also across various subsets of the lexicon, including parts of speech (noun, adjective, or verb), frequency strata (zero vs. nonzero in Francis & Kučera 1982), morphological complexities (prefixed vs. not), and etymological origins (Germanic, Romance, or other).

This section corroborates and in several respects extends Kelly's findings using a different (larger) corpus (CELEX, Baayen et al. 1993) and more thorough controls. Subsequent sections address other languages and phenomena, experimental results, and analysis. Figure 1 illustrates the correlation between initial onset size (in segments) and stress propensity (percentage of words initially stressed in each condition) in a subset of the English lexicon, namely, morphologically simplex disyllables in CELEX (N = 8,323).[1] The ascending line reveals that as onset size increases from zero to three consonants, the incidence of initial stress likewise increases monotonically. Error bars represent 95% confidence intervals (based on Wilson scores for proportions; Wilson 1927, Newcombe 1998, 2000). An asterisk indicates a significant difference between the pair of proportions below it (based on Fisher's exact test with a Holm-Bonferroni correction for multiple comparisons (Holm 1979); all tail-sensitive p-values reported in this article are two-tailed). Thus, both Ø < C and C < CC are independently significant (and CC < CCC is borderline, with $p = 0.02$ before the correction and $p = 0.06$ after it).

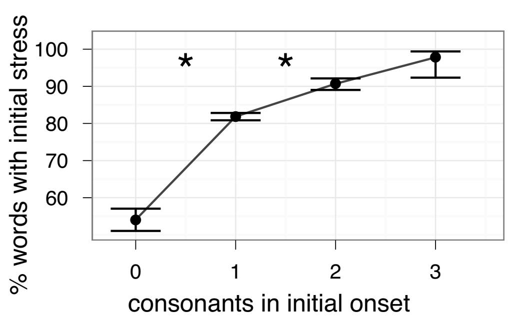

FIGURE 1. Percentages of disyllables in the English lexicon with initial primary stress as a function of initial onset size. Asterisks indicate pairwise significance, with 95% confidence intervals provided.

As Figure 2 shows, the correlation persists across major subdivisions of the lexicon. The first row divides words by part of speech (noun, adjective, or verb). The second depicts near-equally populated terciles according to COBUILD (1987) frequency. The third controls for the skeletal structure of the initial rime, exemplifying the three most frequent types (V̆, V̆C, and VV, where V̆ represents a short vowel and VV a long or diphthongal vowel). This third row emphasizes that the effect is not being driven covertly by the rime, as could be the case if longer onsets tended to cooccur with heavier rimes in the lexicon. Even if consideration is confined to disyllables in which the CV structures of both rimes are held at their modes (i.e. word shape $C_0\breve{V}.C_0\breve{V}C$, N = 1,405), the initial onset effect remains significant (Ø < C and C < CC both with corrected $p < 0.01$).

Although Kelly (2004) addresses only word-initial onsets, onset size and stress also correlate word-medially. Figure 3 is organized like Fig. 2, except that null onsets are omitted, as they are rare (or arguable) medially. The locations of syllable boundaries follow CELEX.

[1] More specifically, words coded in CELEX as complex (C) or contracted (F) were removed, while those coded as monomorphemic (M), zero derivation (Z), 'may include root' (R, e.g. *imprimatur*), or morphology irrelevant (I), obscure (O), or undetermined (U) were retained (see Burnage 1990:§3.1.4).

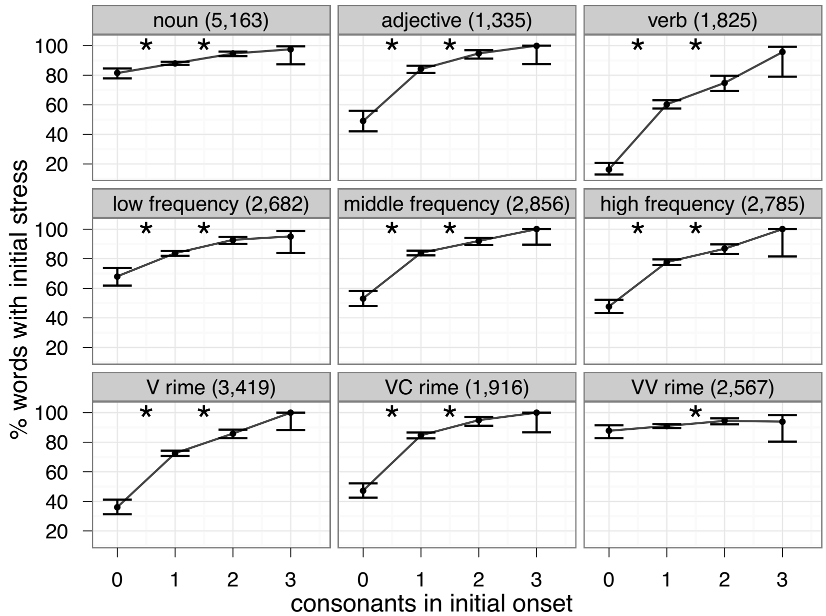

FIGURE 2. Correlations between initial onset size and stress within various subsets of the English lexicon, as labeled, with the population of each subset given (cf. Fig. 1).

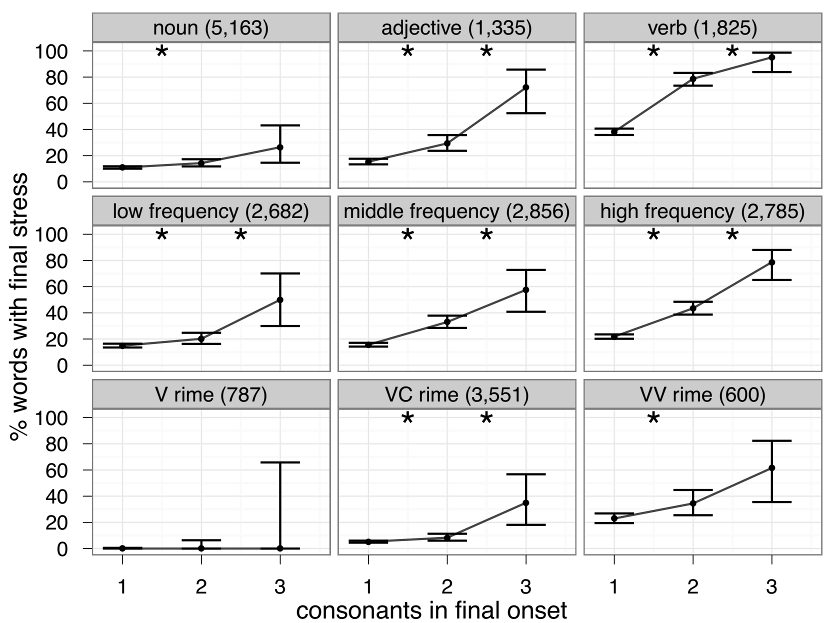

FIGURE 3. Percentages of finally stressed disyllables as a function of final onset size.

These sorts of factors can be combined into a logistic regression model predicting initial primary stress in disyllables. Predictors here include initial onset size (zero to three), final onset size (one to three), initial coda size (zero to two), final coda size (zero to three), initial vowel identity (twenty-three levels), final vowel identity (twenty-four levels), CELEX part of speech (nine levels), and log frequency plus one. This model is more fine-grained than the plots; for instance, it controls for vowel quality. All eight factors are significant ($p < 0.0001$ in an ANOVA). The initial onset coefficient is positive (initial-stress-preferring), with all three of (Ø < C), (C < CC),[2] and CC < CCC being independently significant (Tukey's HSD $p < 0.0001$ for the first two and $p < 0.05$ for the last; note that Tukey's test builds in penalties for multiple comparisons). The final onset coefficient is negative (final-stress-preferring), with both C vs. CC and CC vs. CCC significant ($p < 0.0001$), the latter in each case favoring final stress. Unsurprisingly, in both syllables, heavier rimes are more stress-attracting, where 'heavier' can be understood as possessing more coda consonants and/or a longer vowel. But the model significantly underperforms if the onset factors are removed ($F(2) = 52.5$, $p < 0.0001$).[3]

Figure 4 compares the efficacy of onset vs. rime structure in predicting primary stress placement in simplex disyllables, now crossing initial and final positions with two parts of speech (noun and verb) in the plots. Among rimes, the universal V̆ < V̆C < VV < VVC hierarchy (Ryan 2011a:414) is evident, especially in final syllables. Perhaps surprisingly, judging by the vertical ranges of the lines, onsets appear to be nearly as impactful as rimes. Nevertheless, the visualization may be misleading due to ceiling and floor effects (e.g. the difference between 0% and 5% for final rime V̆ < V̆C in nouns is considerably more significant than the visually more prominent distinction between 13% and 31% for final onset C < CC). Second, the plot is based only on disyllables, while onsets might be less relevant in longer words (next paragraph). Finally, unlike the regressions, the plot does not correct for any correlations between onset and rime structure. Indeed, in §5.1, a comparison of coefficients from logistic regression suggests that onset size is only 46% as effective as coda size in predicting primary stress in disyllabic nouns.

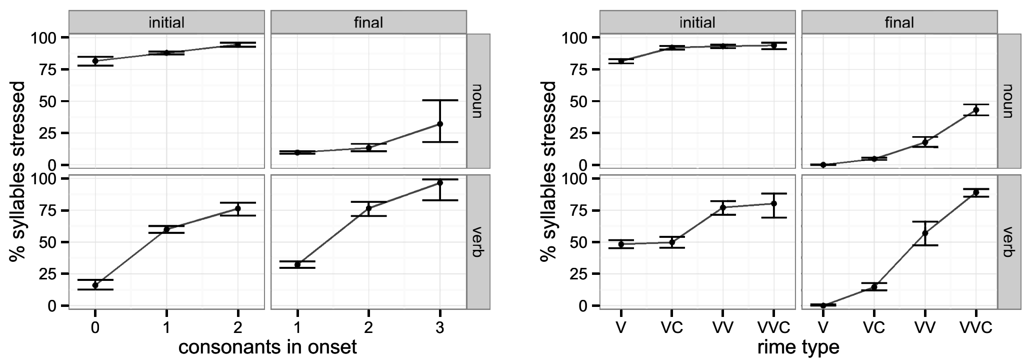

FIGURE 4. Onset and rime conditions compared as predictors of primary stress placement in disyllables. Position in the word (columns) is crossed with part of speech (noun vs. verb; rows).

A logistic model with the same predictors was also trained on simplex trisyllables ($N = 4{,}261$) to test whether the onset effect extends to longer words, which Kelly (2004) did not consider. For the initial onset, Ø < C remains significant (Tukey's $p < 0.0001$), while C vs. CC is nonsignificant (CC vs. CCC, for its part, is too marginal to gauge, with only thirteen CCC items). The second onset contrasts are also nonsignificant. Note, however, that trisyllables are half as frequent as disyllables, so significance tests

[2] In this article, parentheses are sometimes employed in listing multiple contrasts to clarify grouping, since strings like 'Ø < C, C < CC' can be confusing to read.

[3] Since onset-final glides {w, j} are vowel-like, this model was also tested on the subset of data with glide-final onsets removed. In this 5.6% smaller corpus, the onset effects only become stronger, with $F(2)$ (as reported above) climbing to 56.3.

are generally less probative, particularly given the large number of controls. Moreover, syllables tend to be more compressed in longer words (White 2002).

The gradient durations of vowels are also not a confound. If a given vowel tended to be longer following a longer onset, one might argue that vowel length was driving the apparent onset effect. In fact, the opposite correlation is found: onset and vowel durations tend to exhibit a trading (compensatory) relationship. Figure 5 plots the correlations between vowel duration and the number of consonants in the preceding onset for each of thirteen vowels annotated in the Buckeye Corpus of Conversational Speech (Pitt et al. 2007), considering only open, stressed, word-initial syllables of polysyllabic nouns.[4] All thirteen correlations are negative, meaning that the duration of a given vowel tends to decrease as its onset size increases. This anticorrelation has been analyzed by the 'C-center effect' (Browman & Goldstein 1988, Katz 2010:17). If stressability were driven by rime properties alone, one would expect to find either no correlation or a negative correlation with onset size, but not the positive correlation observed in the figures above.

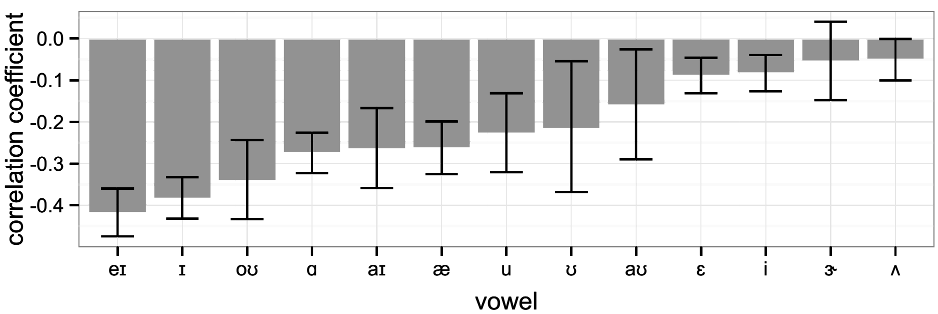

FIGURE 5. Pearson's correlation between vowel duration and preceding onset size for each of thirteen English vowels. Error bars are 95% confidence intervals based on the *t*-statistic.

Finally, though this section employs complexity (segment count) as a proxy for onset size, phonetic metrics such as duration equally demonstrate the correlation. Figure 6 plots the word-initial onset-stress correlation in the subcorpus of disyllabic nouns and adjectives (both of which favor initial stress, unlike verbs and adverbs). The left plot interprets onset size in terms of segment count, the right in terms of mean duration in Buckeye (based on initial, stressed syllables of polysyllabic nouns and adjectives). To reduce clutter, onsets attested fewer than five times in Buckeye or CELEX are excluded, as are glide-final onsets (see n. 3), and font size is proportional to log frequency. The correlation between onset complexity and initial stress (left plot) is r = .685, slightly worse than that between duration and stress (right plot) at r = .722. Even controlling for complexity, duration adds some explanatory power (see also §5). For example, rhotic-final CC onsets tend to be both shorter in duration (mean 153 ms) and less stress-attracting (93% initially stressed) than all other CC onsets (173 ms and 99%, respectively).

In similar fashion, features of onsets, not just complexity, can be tested. Consider, for instance, the typological generalization that a weight criterion distinguishing between voiced and voiceless stops in the onset always treats the latter as the heavier (§1). This

[4] While fourteen stressed vowels are found in Buckeye, one of them, [ɔɪ], occurs in the aforementioned context only following simplex onsets in the corpus. It is therefore omitted in the figure.

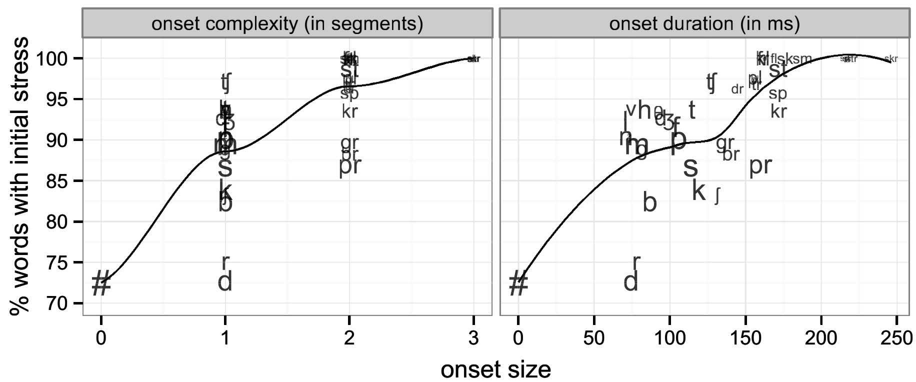

FIGURE 6. Initial stress propensity in disyllabic nouns and adjectives (y-axis) as a function of initial onset size (x-axis). On the left, onset size is discrete (segment count); on the right, it is continuous (mean duration). Font size is proportional to log frequency. # refers to the null onset.

generalization, it turns out, extends to English, albeit statistically rather than categorically. The regression for disyllables above is now repeated, except substituting initial onset voicing for initial onset size and removing all items not beginning with a simple voiced or voiceless stop onset. The voiceless condition is significantly more stress-attracting ($p < 0.0001$).

In sum, the tests in this section broadly support a significant positive correlation between onset size and stress propensity in the English lexicon. The correlation is robust across a variety of subsets of the lexicon, including parts of speech and frequency strata. It is shown not to be confounded by characteristics of the rime (including vowel quality and duration), to hold independently of initial and medial syllables, and to hold independently of disyllables and (to a lesser extent) trisyllables. The effect is monotonic, in the sense that both Ø < C and C < CC (and CC < CCC in some tests) are significant word-initially, and both C < CC and CC < CCC are significant medially. The effect is even subsegmental, with voiceless stops attracting stress more than voiced ones, and CC clusters (where C2 ≠ [ɹ]) more than C[ɹ] clusters, the former in both cases being the longer.

To be sure, a correlation between stress and onset length does not in itself guarantee that longer onsets tend to attract stress. For one, the causality could logically go in the other direction: stressed syllables might tend to license greater complexity (as, for example, in a stage of acquisition of Canadian French in which the two children studied by Rose (2000:133) simplified onset clusters in unstressed syllables but preserved them in stressed syllables; more generally, marked structure is often licensed only in prominent environments (Beckman 1998)). The hypothesis pursued in this article, however, is that longer onsets favor stress because they contribute to syllable weight. A few facts might be cited at this point in favor of this position. First, the behavior of empty onsets is problematic for the markedness-based account. Since a CV syllable is less marked than a V syllable, if stress is merely taken to license greater markedness, the avoidance of stress on onsetless syllables is unpredicted. Second, as described above, an initial CV syllable in which C is a voiceless plosive is more stress-attracting than one in which C is voiced. The lexical statistics of English (this section) align with the categorical stress typology (§1) on this point (on its weight-based explanation, see §5). Nevertheless, it is the more stress-rejecting voiced plosive that is the more marked in this context (Iverson 1983, Westbury & Keating 1986, Fougeron & Keating 1997, Kiparsky 2006). Third, the weight contrast between null and simple onsets (Ø < C) is greater than that between simple and complex onsets (C < CC; see e.g. Fig. 1), even though the markedness differential is likely smaller for the first contrast (see §5.3).[5] Additional arguments for the weight-based position are put forth in the following sections. For one, under wug-testing in which onsets are given but stress is not, the correlations described here persist (§3). Similarly, poetic behavior, in which fixed linguistic material is aligned to metrical templates, supports the same interpretation (§4).

2.2. Russian (and a note on italian). The correlation between onset size and stress propensity established in §2.1 for English is also pervasive in the Russian lexicon. Stress/accent placement in Russian roots is not fully predictable and is therefore assumed to be lexically specified, at least in nondefault cases (Jakobson 1948, Halle 1973a, 1997, Melvold 1989, Revithiadou 1999, Alderete 2001, Crosswhite et al. 2003, Gouskova & Roon 2009). Cubberley (2002:67), for instance, points to the existence of at least 150 minimal pairs for stress among nouns (e.g. *múka* 'torment', *muká* 'flour'). While most analyses of Russian stress therefore take lexical accentuation as a given and focus instead on the rich interplay of stress and morphology, the present treatment considers the predictability of accent within roots.

A corpus of Russian nouns, adjectives, and verbs was derived from a 32,616-lemma frequency list compiled by Serge Sharoff (accessed at http://www.artint.ru/projects/frqlist.php, December 2011; see Sharoff 2002). This list indicates normalized frequency and part of speech. The location of stress, not being indicated in Sharoff's list, was supplied from an online Russian dictionary indicating stress (accessed at http://starling.rinet.ru/cgi-bin/morph.cgi, December 2011). Lemmata for which stress look-up failed were excluded, as were monosyllables, other parts of speech, and compounds in which secondary stress was indicated (cf. Yoo 1992, Gouskova 2010). Lemmata with mobile stress were also removed.[6] After these exclusions, the resulting corpus comprises 24,414 lemmata (11,757 nouns, 5,258 adjectives, and 7,399 verbs).

Figure 7 reveals that the same trends observed for English in Fig. 2 also obtain in Russian. Unlike Fig. 2, however, this figure is based on trisyllables, which are more frequent than disyllables in the Russian data (N = 9,221). The set-up of the third row is also different. It now represents the first intervocalic interlude (C*V<u>C*</u>V…) rather than the first rime (C*<u>VC*</u>.CV…). This adjustment has two motivations. First, Russian lacks phonemic vowel length, so the skeletal structure of the nucleus is moot. Second, the corpus data did not come syllabified, and the proper syllabification of complex interludes in Russian is often unclear (Chew 2003). That said, when the interlude is C alone, the syllabification is uncontroversial (C*V.CV…); in this condition, then, the CV structure of the first rime is held constant, and the effect remains clear. As always, onset size is reckoned phonologically rather than orthographically (e.g. uniliteral щ = {[ɕː] or

[5] Indeed, connecting weight to markedness is arguably not an objection to weight at all, since, for one, almost all canonical, rime-based criteria could also be described in such terms. In Latin, for instance, a heavy syllable is one containing a branching rime (coda and/or long vowel), which is more marked than a nonbranching rime (short vowel with no coda). Of course, stress is not required to license codas or long vowels in Latin, but the same is equally true for English, in which (stress-attracting) voiceless and complex onsets can occur without stress.

[6] That is, only lemmata for which the location of stress is fixed throughout the paradigm were retained. As a heuristic, the following slots of the paradigm were checked: for nouns, the nominative singular, nominative plural, and accusative singular; for adjectives, the masculine nominative singular, masculine short form, and feminine short form; and for verbs, the infinitive, all four preterites, and the first-person singular and third-person plural presents. The vast majority of all three parts of speech are invariant (97.7% of nouns, 95.3% of adjectives, and 78.5% of verbs).

[ɕt͡ɕ]} is two segments, while biliteral пь = [p^j] is one). Yers (ь and ъ) and glides (the onglides of certain vowels, e.g. я [ja]) are never counted as consonants. Onsets of more than three consonants are possible, but rare (0.2% of lemmata), and so are collapsed with three in this section.

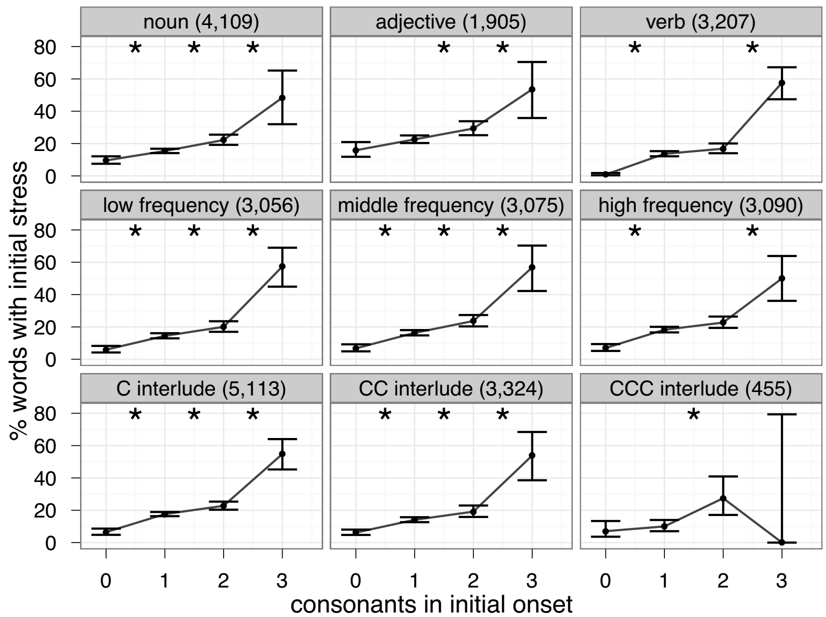

Figure 7. Percentages of initially stressed trisyllables in Russian as a function of initial onset size in various conditions (cf. Fig. 2).

Logistic regression, as in §2.1, corroborates the visualization. The outcome is initial stress in a trisyllable. The predictors are initial onset size (zero to three), initial vowel identity (nine levels), initial interlude length (zero to five), part of speech (three levels), and log frequency plus one. The onset effect is robust: (Ø < C), (C < CC), and CC < CCC are all highly significant (Tukey's HSD $p < 0.0001$), and the model without the onset factor fares significantly worse ($F(1) = 91.5$, $p < 0.0001$). The correlation also extends to words of other lengths (not shown), including disyllables (Ø < C and C < CC both $p < 0.0001$; CC vs. CCC nonsignificant) and tetrasyllables (Ø < C and CC < CCC both $p < 0.0001$; C vs. CC nonsignificant).[7] Unlike English (§2.1), onset stop voicing exhibits no correlation with accent in the Russian lexicon, judging by trisyllables ($p = 0.82$); however, the contrast in 'voicing' is realized differently in the two languages (Petrova et al. 2006).

Although yers (formerly reduced vowels now treated as hard and soft signs in Russian orthography) were not counted above, in order to emphasize that the onset effect is independent of (former) yers, the trisyllable regression was rerun with all (941) yer-containing roots excluded. All three onset contrasts remain significant ($p < 0.0001$) in this 10% smaller corpus. Even if every root with a yer or palatalized consonant is re-

[7] Handbooks addressing historical Russian accentology point to at least one context in which onset size is claimed to condition accent shift (Marek Majer, p.c.): in the nineteenth century, accent shifted from the preverb to the root of a preterite class C verb iff the root began with a complex onset, for example, *só-bral* > *so-brál* but *zá-pil* > *zá-pil* (Vinogradov et al. 1960:480, Garde 1976:276). The generalization may be spurious.

moved, reducing the trisyllable corpus by 51%, Ø < C and CC < CCC remain significant ($p < 0.0001$).

In conclusion, in Russian, as in English, onset size and stress/accent are significantly correlated, not only in the aggregate, but also consistently across various independent subsets of the lexicon, including parts of speech, frequency strata, and following interlude structures. The effects are found independently in lemmata of two, three, and four syllables and are argued not to be confounded by vowel qualities, yers, or palatals. Beyond onset complexity, Gouskova and Roon (2013) find that the quality of segments in complex onsets affects the distribution of secondary stress in Russian compounds (e.g. falling-sonority *ld* is more stress-attracting than rising-sonority *zl*; the extent to which these differences correlate with duration or p-center differences remains to be explored).

As an addendum to these sections on English and Russian stress, Italian stress, though not treated here, is apparently also gradiently onset-sensitive. Hayes (2012) explores a variety of predictors of stress placement and finds that constraints favoring stress on the penult if its onset is complex are assigned nonnegligible weights (see also Davis 1988).

3. Productivity. The onset-stress correlations identified for English and Russian in §2 were based on the lexicon. A pattern, however, might be significant in the lexicon but unproductive (as evidenced by a lack of extension to novel forms), in which case the synchronic grammar would not need to countenance it (cf. Albright & Hayes 2006, Becker 2008, Becker et al. 2011, Hayes et al. 2009, Becker et al. 2012, Hayes & White 2013, though note that the effects in §2 differ from typical accidental generalizations in being consistently monotonic-increasing across contexts). This section addresses the productivity of the onset effect in stress. Further evidence for productivity is adduced from poetic behavior in §4.

Here the groundwork was already laid by Kelly (2004:237), who found initial onset C < CC to be a significant predictor of stress placement in disyllabic nominal pseudowords read aloud from orthographic prompts. For example, *brontoon* was more likely to be initially stressed than *bontoon*. As this pair (one of forty-six such pairs) illustrates, the completion (remainder of the word) was controlled, such that its baseline propensity for initial stress is irrelevant; only the departure from that baseline as a function of onset size was tested. Onset conditions were nested (e.g. *br* contains *b*) and each participant saw only one condition per completion, leaving completion baselines to be factored out in analysis. Ryan (2011b:175) corroborated Kelly (2004) with a different methodology in which online participants self-reported stress judgments for orthographic prompts. Though primarily concerned with the rime, the study also found onset C < CC to be significant.

This section builds on this groundwork in two respects. First, the orthographic prompts of previous experiments raise the possibility of visual confounds, for example, that participants might tend to stress larger visual syllables (cf. Seva et al. 2009 and Arciuli et al. 2010 on print-to-speech translation, Colé et al. 1999 on the visual syllable). The two experiments in §3.1 address this concern. Second, §3.2 considers the role of analogy.

3.1. Two experiments controlling for possible visual confounds. To deconfound possible visual interference, a wug test (Berko 1958) was run along the lines of Kelly 2004 and Ryan 2011b above, except using auditory rather than orthographic prompts (on wug-testing stress auditorily, see Guion et al. 2003, Shelton 2007, and Carpenter 2010). The experiment was conducted online using Amazon's Mechanical Turk (see Daland et al. 2011:203). Participants were screened for US location, prior approval of at least 95% on $N \geq 50$ tasks, informed consent, and self-reported native proficiency in English (to disincentivize exaggeration, participants were informed—truthfully—that they would be paid regardless of their proficiency). Each participant was paid $0.50 for a roughly two-minute task. Participants were instructed that in order to help hone a text-to-speech synthesis system, they would listen to sixteen audio clips of words, indicate which syllable sounded more stressed (or 'emphasized') by pressing a radio button, and transcribe the word however they saw fit. Some of the words, they were informed, would be real, others made up.

The sixteen audio prompts were divided evenly between pseudowords (test items) and real English words (fillers). The order of all items (test and filler) was randomized across participants, except that the first two items were always fillers and no two subsequent fillers were ever adjacent. The eight fillers, 50% with initial stress, were always {*bamboo*, *gazelle*, *giraffe*, *machine*, *pamphlet*, *railroad*, *redwood*, *sawdust*}. The eight test items are given in orthography in Table 1. Each participant was exposed to one onset condition per item.

	Ø CONDITION	C CONDITION	CC CONDITION
1.	ummorm	**l**ummorm	**fl**ummorm
2.	izzoof	**r**izzoof	**gr**izzoof
3.	irgeen	**l**irgeen	**fl**irgeen
4.	illawm	**w**illawm	**sw**illawm
5.	izzool	**b**izzool	**br**izzool
6.	ordoot	**m**ordoot	**sm**ordoot
7.	evvain	**d**evvain	**dr**evvain
8.	istrow	**l**istrow	**sl**istrow

TABLE 1. Eight test items × three onset conditions for an auditory experiment.

All items and fillers were recorded by a male speaker (mono 44.1 kHz) with final stress and manipulated in Praat (Boersma & Weenink 2011) to normalize pitch, intensity, and duration. Pitch was flattened to 150 Hz and intensity to 65 dB, rendering the clips rather unnatural (recall that participants were prepared to hear synthesized speech). For test items, nonnull onsets were spliced onto the null-onset completion to hold the completion identical across conditions.[8] This can be seen in Figure 8 for the final row of items in Table 1. Fillers (real English words) were pitched up slightly (5 Hz) on their natural stresses, in part to encourage participants to continue to listen for subtle stress differences. Each item was isolated in its own file with an obfuscated filename, padded with 500 ms of silence on both sides, converted to MP3, and embedded in the experiment using the Google Reader audio player.

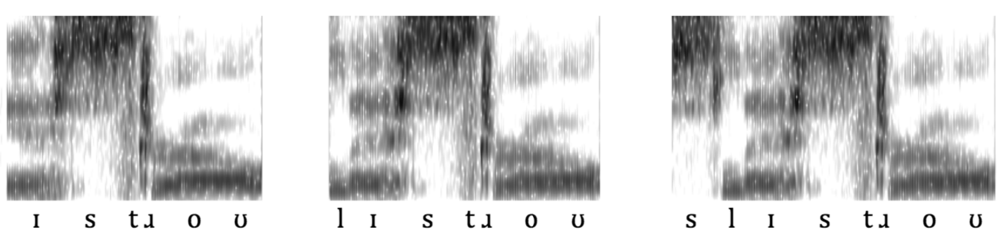

FIGURE 8. Spectrograms of *istrow*, *listrow*, and *slistrow* showing phonetically identical completions.

[8] A fixed portion of the initial vowel (150 ms) was also overwritten to retain natural-sounding transitions without altering the pitch, intensity, or duration of the completion.

In addition to the screening criteria above, participants were analyzed only if they answered correctly for at least seven of the eight real English words, ensuring understanding of the task and functionality of the interface. Thirty-eight of 166 participants met these criteria. The onset factor is significant as a predictor of initial vs. final stress ($F(2) = 12.7$, $p < 0.0001$), with 43% initial stress for Ø, 60% for C, and 79% for CC. Under logistic regression with random effects for participant and completion, Ø < C and C < CC are independently significant (both Tukey's HSD $p < 0.05$). (See also Fig. 10 below for more detailed results.) Thus, the onset effect crosscuts visual and auditory modes of presentation.

A second experiment corrects for visual confounds by taking advantage of digraphs. Participants, recruited as above, pressed radio buttons indicating their stress preferences for sixteen items, including six test items (wugs) and ten fillers (real English words). Only participants scoring at least nine out of ten on the fillers, in addition to the other criteria above, were analyzed. Thirty-six of eighty-four participants met the criteria. All items were presented orthographically in the frame 'A ___ is a kind of ___ ', where the second gap was filled by an appropriate classifier for real words and by one randomly selected from a set (*fish*, *fabric*, *axe*, *flower*, etc.) for wugs. Each participant saw each test item in either a simple (digraph) or complex condition; see Table 2. The number of letters is thus the same in both conditions. Test items and fillers were randomly interspersed as described above.

	C (DIGRAPH) CONDITION	CC CONDITION
1.	**th**appeen	**dr**appeen
2.	**ph**olmank	**st**olmank
3.	**sh**areem	**st**areem
4.	**sh**evveen	**sm**evveen
5.	**sh**illect	**br**illect
6.	**th**izzarn	**gl**izzarn

TABLE 2. Six test items × two onset conditions for an orthographic experiment.

The rate of initial stress for digraph-initial pseudowords was 40%, significantly less than the 56% rate for the same completions with complex onsets (Fisher's exact test $p = 0.03$). The effect remains significant ($p < 0.05$) under logistic regression with random effects for participant and completion. Five of the six items in Table 2 received initial stress less often in their digraph conditions, the one exception (item 1) being far from significant ($p = 0.65$). In sum, when letters per visual syllable is controlled in orthographic stimuli, the onset effect persists. All protocols attempted thus far—pronunciation of orthographic wugs (Kelly 2004), stress judgments of orthographic wugs (Ryan 2011b), stress judgments of auditory wugs (here), and stress judgments of visually controlled orthographic wugs (here)—converge in support of a productive onset effect.

3.2. ANALOGY. This section addresses the question of whether the extension of the onset effect to pseudowords could be driven entirely by analogy, obviating the need for a grammatical principle affecting onset weight. One could imagine a scenario under which syllables with longer onsets tended to accrue stress diachronically (perhaps owing to misperception, though this would still require an explanation) without any synchronic constraint favoring the association (cf. Blevins 2004). Under this scenario, one might attribute the striking consistency of the pattern in the lexicon to channel bias (cf. Moreton 2008, Yu 2011, Sonderegger & Niyogi 2013) and its projection onto novel items to analogical induction. This section argues that analogy alone is unlikely to suffice. The onset effect emerges even in neighborhoods in which it is locally unsupported (or reversed), suggesting broad grammatical generalization.

An analogical model projects a lexical neighborhood for a novel item according to some similarity metric, with neighbors then voting on its treatment. Depending on the model, the neighborhood might be as small as the single most similar item or as large as the entire lexicon, and neighbors might vote equally or unequally (their weights a function of similarity). Figure 9 illustrates these principles for a novel English noun, *plizzoof* /plɪzuf/. The neighborhood shown is approximately 150 items, all disyllabic nouns, with the most similar items (printed largest and closest to the origin) holding greatest sway. The figure is divided into two panels for iambic (left) vs. trochaic (right) neighbors. As suggested by the total masses (in arbitrary units) given above the panels, final stress is predicted to be favored for *plizzoof*.

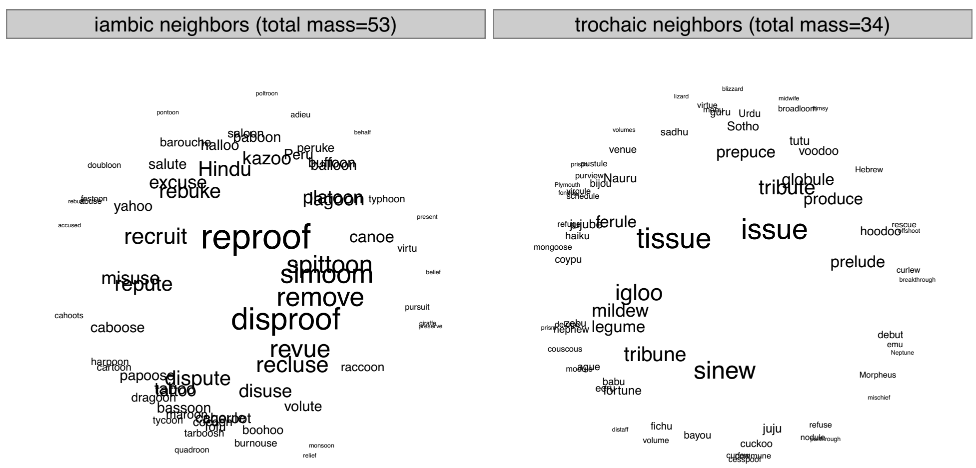

FIGURE 9. Schematic illustration of neighbors of the pseudoword *plizzoof*. Only disyllabic nouns are shown, in orthography. Similarity, and hence influence, is represented by both font size and proximity to the origin. θ-coordinates are arbitrary.

Two well-known analogical platforms are tested here, namely, ANALOGICAL MODELING (AM; Skousen 1989, 1992, 2009, Eddington 2000, Skousen et al. 2002; cf. Daelemans et al. 1994) and the TILBURG MEMORY-BASED LEARNER (TiMBL; Daelemans et al. 2010; see also Daelemans & van den Bosch 2005 and Skousen et al. 2002:Part IV). Since the wug tests described in §3.1 involve only disyllabic nouns, the models in this section are likewise trained on the subcorpus of simplex disyllabic nouns in CELEX (§2.1), though this is an oversimplification (a charitable one, since it preselects features known to be relevant, obviating the models' need to ascertain their relevance). As another point of supervision, words are given to the analogical models with subsyllabic structure, permitting the alignment of onsets, nuclei, and codas across items; the models perform substantially worse if this structure is not provided.

For existing disyllabic nouns, AM exhibits leave-one-out classification accuracy of 93.5% (for comparison, a model assigning uniform initial stress is 87.0% accurate). TiMBL is tested here under a range of *k*-nearest neighbors rubrics to locate an optimal model for stress (for a similar winnowing of TiMBL models, see Hayes et al. 2009:855). A space of seventy-two models was searched, representing every permutation of the parameters metric ∈ {overlap, MVDM, Levenshtein, Dice coefficient}, feature weighting ∈ {gain ratio, information gain}, *k* neighbors ∈ {1, 3, 7, 11, 19}, and (if

$k > 1$) neighbor weighting ∈ {equal, inverse linear} (see Daelemans et al. 2010:39ff. for details). A maximum accuracy of 94.3% was achieved with the specifications <overlap, information gain, 7, inverse linear>. This best model is therefore taken to represent TiMBL below.

Having been trained on existing disyllabic nouns, the models can then generate predictions for pseudowords. Figure 10 compares the observed rates of initial stress for the first experiment in §3.1 to the predictions of AM and TiMBL, paneled by completion, with the rightmost plot showing the means over completions. As the slopes of these means suggest, AM predicts an aggregate onset effect; TiMBL does not. But closer inspection reveals the observed onset effect to be more consistent across completions than either model predicts, emerging as a positive correlation even when the models predict it to be negative. The observed incidence of initial stress increases in 95% of the sixteen adjacent comparisons in Fig. 10, but analogy predicts it to increase in only 56% (AM) or 50% (TiMBL) of comparisons. Analogy is undergeneralizing.

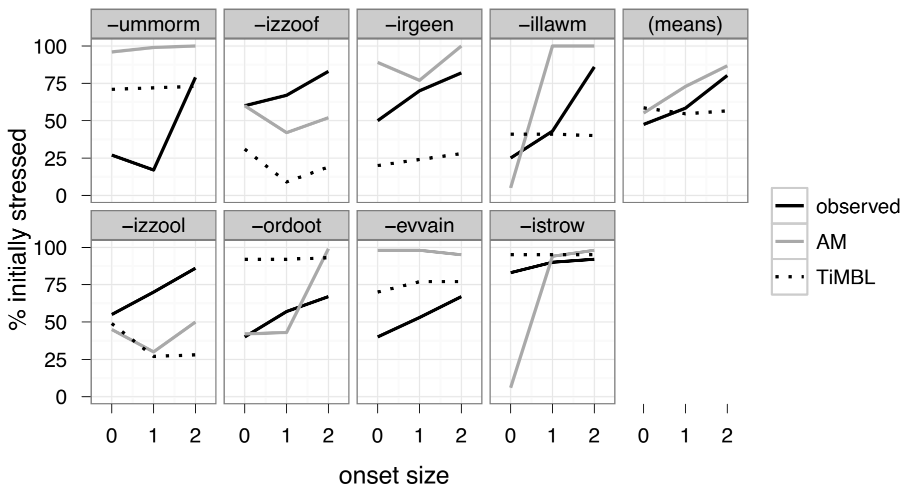

FIGURE 10. Correlation of onset size and stress in experimental results (solid black lines) alongside predictions of two analogical models. The rightmost plot shows means across completions.

The inadequacy of the analogical models can be demonstrated more rigorously with logistic regression. A model with analogical predictors alone (one factor representing AM, another TiMBL) significantly underperforms the superset model with an added factor for onset size ($F(1) = 17.5$, $p < 0.001$). Applying the same diagnostics to the second (digraph) experiment in §3.1, the onset effect is likewise found to be significant above and beyond both analogical models ($F(1) = 6.8$, $p = 0.01$).[9] Thus, while analogy can replicate the onset effect once it is established and locally supported, experimental data suggest that the effect is more general, emerging also when unsupported or even negated by local comparanda. This fact, coupled with the consistency of the correlation

[9] The experimental results reported by Kelly (2004:238, 243) also appear to overreach these models, though significance cannot be tested here, since the complete data are not available. Kelly tested forty-six items in C and CC conditions, finding thirty-six to be initially stressed more often in the CC condition, four to be ties, and six to reverse the trend. For the same items, AM predicts thirty-three increases, three ties, and ten reversals, and TiMBL twenty-four increases, ten ties, and twelve reversals, both models doubling or nearly doubling the observed frequency of reversals.

across lexicons (§2) and metrical corpora (§4), which analogy cannot motivate, supports the grammaticality of the effect.[10]

4. POETIC EVIDENCE FOR PRODUCTIVITY. In syllabic (as opposed to mora-counting) quantitative meters, metrical positions vary in their tolerances for heavy vs. light syllables. Some such meters are rigid, in that a position tolerates only syllables of a particular weight category, with no exceptions (e.g. Deo 2007). Others are less rigid, in that positions exhibit flexible preferences. Ryan 2011a showed that gradient continua of syllable weight can be extracted from flexible meters. For example, if a poet places a heavy syllable in a preferentially light position, it will tend to be a lighter heavy more often than chance (or a baseline from other positions) would predict, permitting the inference of a weight continuum from the relative proportions. Ryan 2011a considered only properties of the rime. This section demonstrates that the poets' placement of syllables is also significantly sensitive to the onset. In particular, syllables with longer onsets are underrepresented in preferentially light positions, all else being equal.

4.1. VEDIC SANSKRIT. The Rig-Veda (c. 1200 BC) is the oldest extant Sanskrit (or, more properly, Vedic) text. The edition employed here, slightly updated from van Nooten & Holland 1994, contains 39,833 lines. The vast majority (97%) of these lines instantiate one of three metrical types, namely, *gāyatrī* (38% of the text), *triṣṭubh* (42%), or *jagatī* (17%), being eight, eleven, and twelve syllables, respectively.[11] In addition to being syllable-counting, Vedic meter is flexibly quantitative, in that positions of the line vary in their propensities for light (short-vowel-final) and heavy syllables, accent being irrelevant. Figure 11, for example, shows the percentage heavy in each of the eight positions of the *gāyatrī* line type, which exhibits an iambic cadence.[12]

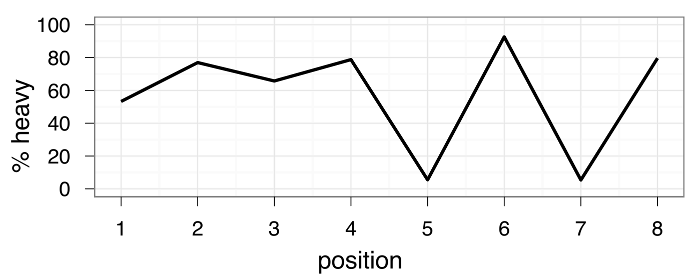

FIGURE 11. Percentage heavy across positions of eight-syllable Vedic lines.

If onsets affect syllable weight, one might expect syllables with longer onsets to be overrepresented in heavier positions, all else being equal (e.g. if one were to compare light syllables in position 6 of the *gāyatrī* to those in position 7, the former might be expected to possess aggregately longer onsets). Nevertheless, the preceding syllable is a confound. For example, a null onset entails vowel hiatus (V.V), a configuration in

[10] This conclusion leaves open the possibility of grammar and analogy interacting in determining stress placement in novel items (Guion et al. 2003).

[11] More properly, these are not meters but *pāda* (line) types, each being the basis of several meters depending on the organization of the stanza. Thus, what is termed *gāyatrī* here comprises all octosyllables, including the *gāyatrī* proper, *anuṣṭubh*, *paṅkti*, and so forth (Oldenberg 1888, Arnold 1905, Macdonell 1916).

[12] Such a plot oversimplifies in two respects, both of which are harmless to the present purposes. First, in showing only unigram propensities, it overlooks certain syntagmatic tendencies. Second, it ignores the heterogeneity of the corpus. For example, a heavy syllable in position 7 is found most frequently in certain subcorpora (e.g. the so-called trochaic *gāyatrī* and epic *anuṣṭubh*), owing not to greater general flexibility but to the development of a particular optional inversion.

which the first vowel normally shortens (Gunkel & Ryan 2011). Thus, a null onset can only follow a light syllable, which in turn is itself attracted to light positions, which tend to be followed by heavy ones. Similar confounds arise for other onset types as well. For example, a complex onset cannot follow a light syllable in Vedic (e.g. /$\breve{V}$#CCV/ is resyllabified as $\breve{V}$#C.CV).

These confounds can be addressed by comparing onset conditions in frames in which the preceding rime is held constant, for example, $C_0\breve{V}\#\underline{VC_0}$ vs. $C_0\breve{V}\#\underline{CVC_0}$ (null vs. simple) and $C_0\breve{V}C\#\underline{CVC_0}$ vs. $C_0\breve{V}C\#\underline{CCVC_0}$ (simple vs. complex). This choice of frames also renders a number of finer points of syllabification irrelevant; for example, resyllabification is moot, since the syllable and word boundaries coincide. A separate linear regression was run for each comparison, taking as data in each case only syllables in the appropriate frame ($C_0\breve{V}\#\underline{C_0^1VC_0}$ for the first test, $C_0\breve{V}C\#\underline{C_1^2VC_0}$ for the second). The outcome is the heaviness propensity (standardized proportion heavy) of the position occupied by the underlined syllable, where propensity is conditioned on meter, since positions are not commensurate between meters. Predictors include the binary onset condition and the skeletal structure of the rime (six levels). This second factor corrects for possible collinearity between onset condition and rime type, given that the rime is known to influence metrical alignment.

In the first comparison, onset Ø patterns as significantly lighter than C, both aggregately ($p < 0.0001$) and in each of the three meters tested independently ($p < 0.01$). Onset C < CC is likewise significant both in the aggregate and in each meter considered separately ($p < 0.0001$).[13] Figure 12 illustrates the consistency of this difference across both meters and rime types, considering the three most frequent conditions of each (eight, eleven, or twelve syllables and $\breve{V}$, $\breve{V}$C, or VV rime, respectively). In every one of the 3 × 3 = 9 comparisons, the CC condition patterns as heavier than the C condition, with the differences being most pronounced among light ($C_0\breve{V}$) syllables. In fact, in the eight-syllable meter, the difference between C$\breve{V}$ and CC$\breve{V}$ is so great that the latter is intermediate between heavy and light. The greater impact of onset size in syllables with lighter rimes is possibly due to a proportionality (Weber's law) effect (cf. e.g. Lunden 2006, 2011).

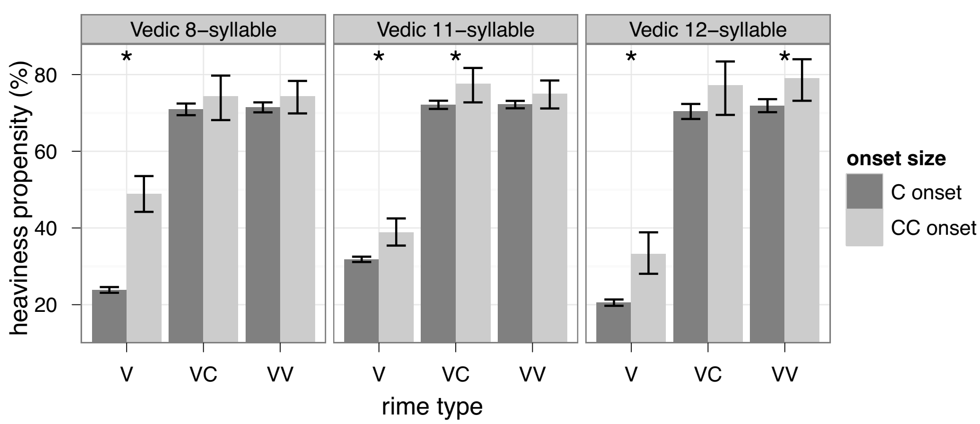

FIGURE 12. Mean heaviness propensities of syllables with CC as opposed to C onsets, grouped first by metrical type and then by rime type.

[13] A referee asks whether there is evidence of quantitative preferences in the precadence, given that the small bumps in the first four positions of Fig. 11 might be due to chance. Though the metrical status of the precadence is not critical here, I note that Ø < C is significant in both the precadence and cadence taken separately ($p < 0.0001$), while C < CC is significant only in the cadence.

In conclusion, onset size correlates monotonically with metrical weight propensity in Vedic meter, not only in the corpus as a whole but also across the three major metrical types considered separately. Moreover, the effect holds across rime structures, further demonstrating its robustness and independence of the rime.

4.2. EPIC SANSKRIT. The two Sanskrit epics, the Mahābhārata and Rāmāyaṇa, both postdate the Rig-Veda by roughly a millennium. This section focuses on the Rāmāyaṇa, a corpus of (here) 38,038 lines (accessed at http://sub.uni-goettingen.de, 2005; Goldman 1990). Of these lines, 95.4% are sixteen syllables long, being of the *śloka* [ɕloːkə] meter; other line types are put aside here.[14] Figure 13 depicts the proportion heavy in each of the sixteen positions. The dashed line after position 8 indicates the fixed caesura.

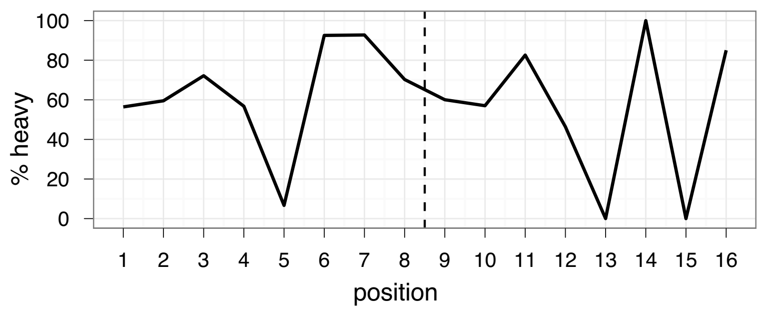

FIGURE 13. Percentage heavy across the sixteen positions of the *śloka* line. The dashed vertical line indicates the caesura.

The two tests described in §4.1 are applied to this new corpus. Both Ø < C and C < CC are highly significant ($p < 0.0001$), meaning that longer onsets are more skewed toward heavier positions, even while controlling for the preceding and following rimes.

4.3. THE FINNISH KALEVALA. The meter of the Finnish Kalevala epic (Lönnrot 1849) is a trochaic tetrameter, that is, four repetitions of strong-weak. Primary stressed syllables, which are always word-initial, must be light (i.e. short-vowel-final) in weak positions and heavy in strong ones (Sadeniemi 1951, Kiparsky 1968, Leino 1994). Mapping is flexible, but stricter toward the end of the line (Figure 14). While the cited descriptions imply that only word-initial syllables are regulated, noninitial syllables weakly shadow the pattern, as shown by the dashed line in Fig. 14, in which 'other' comprises both unstressed syllables and those with secondary stress (excluding $C_0\check{V}$ clitics such as *ja* 'and' from all counts). The solid line does not extend to position 8 because no monosyllables occur there. Similarly, the dashed line does not start with position 1 because the line must begin with either a stressed syllable or a proclitic (excluded here). The present corpus (from http://www.kaapeli.fi/maailma/kalevala, 2009) contains 15,846 octosyllabic lines, excluding lines of other lengths.

Since Kalevala Finnish lacks complex onsets, only Ø vs. C can be tested. A linear regression, set up as in §4.1 and §4.2, reveals that post-$\check{V}$ C is significantly ($p = 0.0007$) more strong-skewed than post-$\check{V}$ Ø, even while correcting for the following rime as before.

[14] The sixteen-syllable *śloka* line is at its root a pair of eight-syllable *pāda*s of the *gāyatrī* (or *anuṣṭubh*; see n. 11) type described in §4.1. Nevertheless, given the systematic discrepancies between half-lines in Fig. 13, it is sensible (and consistent with tradition) to treat the sixteen-syllable unit as the line in this context.

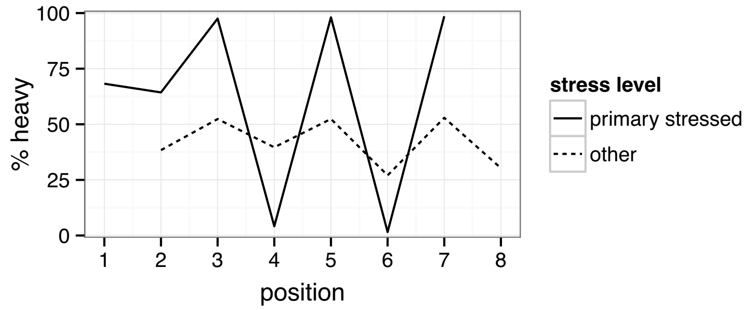

FIGURE 14. Percentage heavy across the eight positions of the Kalevala line.

4.4. ONSET QUALITY AND SUMMARY. While §4.1–4.3 consider only onset complexity, features of onsets can also be probed. The same five metrical corpora (three Vedic meters, Epic Sanskrit, and Kalevala Finnish) are now tested for a contrast between voiced (D) and voiceless (T) stop onsets,[15] the prediction being that if the two significantly differ, the latter will be the heavier (as in §1 and §2.1). As before, the dependent variable is the weight propensity of the position in which a syllable is placed. Predictors include onset voicing and rime structure. Only syllables with simple stop onsets are analyzed. Unlike previous models, the preceding rime is left unconstrained, since the confounds in §4.1 concerning differing onset complexities are moot here. In all five corpora, voiceless stops pattern as significantly heavier than voiced ones ($p < 0.0001$).

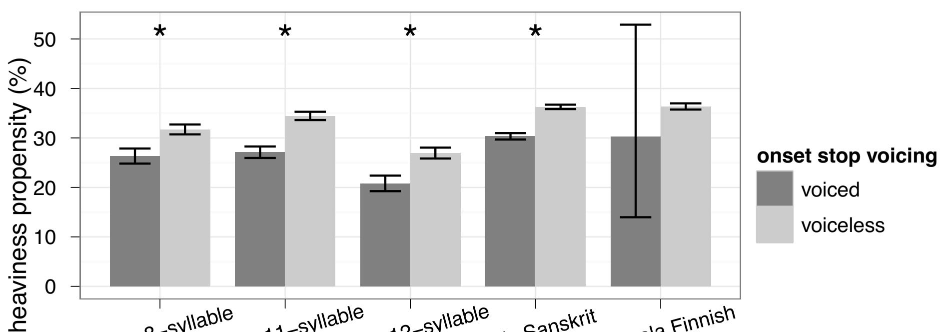

FIGURE 15. Heaviness propensities of DV̆ vs. TV̆ in five meters.

Figure 15 depicts the inferred weights of syllables beginning with voiced and voiceless stops in each corpus, considering only light syllables (DV̆ vs. TV̆) so that the structure of the rime is controlled. While the contrast is nonsignificant in Finnish in the condition shown, it is aggregately significant in the regression, as reported above. Note, however, that voicing is not contrastive in Kalevala Finnish; in fact, the corpus records [ɡ] only as a postnasal allophone of /k/ (and as a further complication, modern pronunciation, at least, renders this string as [ŋː] rather than [ŋɡ]). Regardless of Finnish, the predicted effect is clear and consistent in Indo-Aryan, where voicing is contrastive in the stop series.

[15] See below for a caveat concerning Kalevala Finnish, in which voicing is not contrastive. For Vedic/Sanskrit, only plain voiced and voiceless stops are considered here, aspirated and breathy-voiced stops having been put aside.

LANGUAGE	METER	SYLLABLES PER LINE	Ø < C	C < CC	D < T
Vedic	*gāyatrī*	8	*	**	**
Vedic	*triṣṭubh*	11	**	**	**
Vedic	*jagatī*	12	*	**	**
Sanskrit	*śloka*	16	**	**	**
Finnish	Kalevala	8	**	N/A	(**)

TABLE 3. Summary of onset effects tested in §4.1–4.4 for flexible quantitative meters. One asterisk here indicates significance at $p < 0.01$, two asterisks, $p < 0.0001$. See the text concerning Finnish.

Table 3 summarizes the metrical evidence adduced in this and the preceding sections, which unanimously supports onsets as factors in syllable weight. Since metrics is not the primary concern of this article, it is left to future research to test additional quantitative meters using the sorts of methodologies and controls illustrated here.

5. A RHYTHMIC THEORY OF SYLLABLE WEIGHT. This section turns to the analysis of the onset effects established in §2–4. It is proposed that the span over which the percept of syllable weight is computed begins not with the left edge of the rime, as is usually assumed, but with the p-center (perceptual center; Morton et al. 1976) of the syllable, an event corresponding to the downbeat or perceived beginning of the syllable and serving as the target for isochrony in regularly timed speech (e.g. Patel et al. 1999, Villing et al. 2003, Barbosa et al. 2005, Soraghan et al. 2005, Tilsen 2006, Port 2007, Wright 2008, Villing 2010, and references therein). When a speaker utters syllables in regular succession (e.g. *one, two, three, four*), the left edges of both the syllables and the rimes are systematically anisochronous. For example, if one claps on each syllable, or utters the sequence to a beat (as with rapping), the beats can be seen to align more closely with the beginning of the rime than with the beginning of the onset, though the targets deviate from both. Thus, the p-center is not a syllabically defined event, but a perceptual function whose exact characterization remains unclear (op. cit.). The purpose here is not to contribute to the p-center problem, but to propose a possible line of synthesis between syllable rhythm and syllable weight as at least part of the explanation of both onset weight effects and of the subordination of onsets to codas as contributors to weight.

5.1. ONSET VS. RIME EFFECTS ON P-CENTER LOCATION. As a concrete example, Figure 16 illustrates p-centers for two English monosyllables, *ba* and *spa*, as uttered by the first speaker ('DY') in the Harvard-Haskins Database of Regularly Timed Speech (Patel et al. 1999). Zero on the x-axis represents the beginning of the rime, marked by the onset of periodicity in the waveform. Below each waveform, a density curve (smoothed histogram) depicts the distribution of p-centers for this speaker (normalized as in Patel et al. 1999). Port's (2007:509) description agrees with this (independently derived) figure in that 'for *ba*, the beat occurs right at the onset of the vowel, and for *spa*, the beat moves slightly to the "left" of the vowel onset showing that the [s] in *spa* has some influence on the effective location of phase zero'.

Some p-center research has attempted to gauge the relative impacts of onset vs. rime duration on p-center placement (among other factors, such as the shape of the energy envelope). For instance, Marcus (1981:253) finds that if one regresses on the durations of the onset and rime, the optimal coefficients for predicting p-center offset from the beginning of the syllable are 0.65 times the onset duration plus 0.25 times the rime duration (see also Goedemans 1998:94). If the two onsets in Fig. 16 are taken to be 90 and 155 ms, respectively, and both rimes 155 ms, this formula predicts a 23 ms earlier p-center for *spa* than for *ba*, close to the observed difference of 24 ms (mean offset

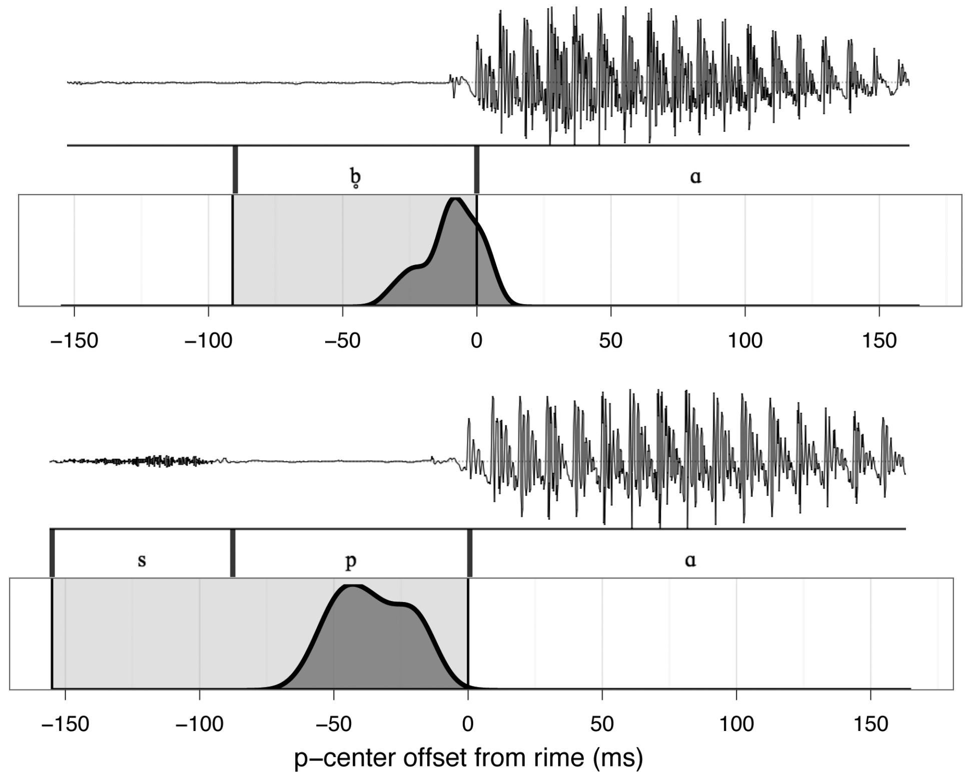

FIGURE 16. P-center densities for *ba* [ˈb̥ɑ] (top) and *spa* [ˈspɑ] (bottom), each below a segmented amplitude waveform illustrating the alignment between acoustic events and p-centers (the left edge of [b̥] is arbitrary).

−4 ms for *ba* and −28 ms for *spa*). In general, then, longer onsets induce earlier p-centers with respect to the rime.[16]

The empirical effect of onset complexity on p-center location is shown more generally in Figure 17. As initial onset size increases from zero to three (top to bottom), p-centers increasingly precede the rime (every step significant with Wilcoxon rank-sum test, $p < 0.01$). Furthermore, though not shown, p-centers occur significantly earlier for voiceless than voiced stops ($W = 1058$, $p < 0.001$).[17] These data are based not on the Harvard-Haskins Database, which lacks Ø and CCC onsets, but on an illustrative corpus of 342 English monosyllables selected randomly from CELEX and uttered by two speakers, one male and one female, to a metronome (88 BPM with syllables on alternating beats). The mean p-center shift (26 ms per added consonant) is only a fraction (35%) of the mean added duration (74 ms per consonant). This 35% rate agrees with Marcus's (1981) coefficient cited above: according to his formula, if one holds the duration of the rime constant, each millisecond added to the onset is predicted to pull the p-center leftward by 35% of a millisecond.[18]

[16] A referee asks whether referring to onset size here and elsewhere should be taken to imply that p-centers are sensitive to syllabification (e.g. different for /VCCV/ syllabified as VC.CV vs. V.CCV). Since only word-initial onsets are considered in this section, the question is moot, and left open, for the present purposes.

[17] Because CCC onsets must contain voiceless stops in English, voicing is partially confounded with complexity here. Testing CCC against the subset of CC with voiceless stops, however, still yields a significant difference ($W = 591$, $p < 0.001$).

[18] The coefficient for onset duration given above, 0.65, represents the p-center offset from the beginning of the syllable. To calculate the leftward displacement of the p-center, one must take its complement, 0.35.

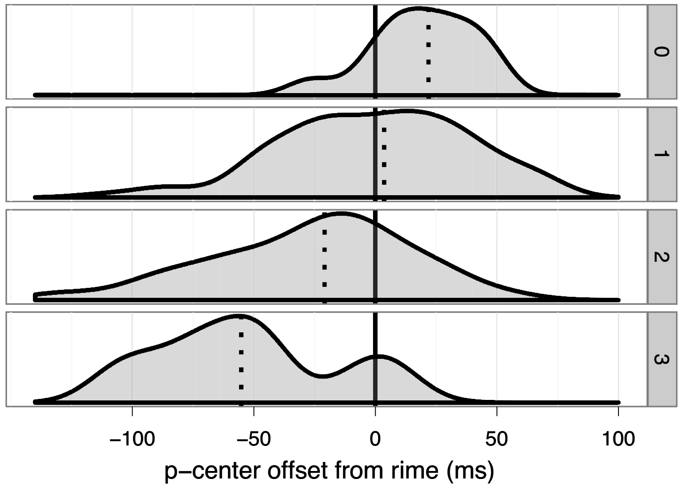

FIGURE 17. P-center densities by onset complexity (from null at top panel to three at bottom). The solid vertical line is the onset-rime boundary, and the dotted line is the median p-center.

It is proposed that the domain of syllable weight begins not with the onset-rime boundary but with the p-center. This move predicts onsets to affect weight, but not to the same extent that codas do, the latter being parsed fully into the weight domain. Based on the p-center data just discussed, the contribution of (each unit of) onset duration to the weight percept is expected to be on average 35% that of (each unit of) coda duration. Marcus's (1981) coefficients predict the somewhat greater value of 47% for this ratio.[19] At any rate, a range of roughly one-third to one-half can be assumed as a working hypothesis. This range (especially the latter benchmark) seems also to jibe with Goedemans's (1998:75) finding that the just noticeable difference (JND) for duration was 47% smaller within the rime (26 ms) than it was within the onset (49 ms), given the same 300 ms comparandum in every case.

We can now return to the gradient weight systems in §2–4 to check whether it is indeed the case that onset consonants affect weight by roughly 35–47% as much as coda consonants do. For example, consider the set of simplex disyllabic nouns in English (upper left panel of Fig. 2). The relative contributions of the onset and coda in predicting stress placement can be gauged by logistic regression, with primary stress as the outcome and onset size, coda size, vowel length, and position in the word as predictors. The onset coefficient (0.45) is 46% as great as the coda coefficient (0.97) in this case. Poetic data can be similarly diagnosed. For example, as Fig. 12 showed for three Vedic meters, the average heaviness propensity of CV̆ is 0.25, while those of CCV̆ and CV̆C are 0.43 and 0.71, respectively. Thus, adding a consonant to the onset increases the propensity by 39% as much as adding one to the coda. Both of these rates (46% and 39%) fall within the hypothesized ballpark of 35–47%.

In sum, while onset consonants contribute to syllable weight, they are underprivileged with respect to coda consonants, an asymmetry predicted by the p-center interval. This span is also supported by the fundamentally rhythmic natures of stress and meter, both phenomena being characterized by regular temporal alternation and hierarchical organization of prominence (Liberman 1975, Liberman & Prince 1977, Hayes 1995,

[19] We are interested here in the interval between the p-center and the end of the syllable as the hypothetical weight domain. Since Marcus's original coefficients predict the p-center offset from the beginning of the syllable, the ratio of their complements (0.35/0.75) is taken to derive the stated value of 47%.

Fabb & Halle 2008). Given the emerging consensus in the experimental literature cited above that linguistic isochrony is not anchored by the beginning of the rime, it is sensible to pursue a unified account of the domains of rhythm and weight outside of the rime. It is worth noting, however, that while stress and meter are treated as rhythmic, it is unclear whether this treatment is equally warranted by all phonological phenomena described as weight-sensitive (see e.g. Topintzi 2010:207 on contour-tone licensing and Gordon 2006 on process-specificity in weight). The p-center proposal may be relevant only for rhythmic weight systems, including stress and meter.

5.2. Constraint-based analysis. As with all phonetic approaches to syllable weight (e.g. Archangeli & Pulleyblank 1994, Hubbard 1994, Broselow et al. 1997, Goedemans 1998, Gordon 2002, 2005, 2006), the weight percept is computed on a phonetic representation. These previous accounts treat the phonetic grounding of categorical distinctions. For example, Gordon (2002) argues that languages tend to select criteria that maximize the dispersion between categories in a perceptually defined space. Thus, even if onsets affect the weight percept, as argued here, their rare recruitment by categorical criteria would follow from their weak effect on the p-center relative to that of the rime. Given the dominance of the rime, a language with codas and a binary criterion is predicted to opt for a coda and/or nucleus criterion over an onset one. In short, the rarity of onset-sensitive categorical criteria is not only unproblematic for the present proposal, but also predicted by it.

At the same time, onset weight is predicted to emerge universally in gradient weight systems, such as those discussed here, in which the heavier a syllable is, the more likely it is to be stressed or to occupy a strong metrical position. Gradient weight can be modeled either by multiplying the number of categories until a fine-enough grain of resolution is achieved or by permitting the grammar direct access to the weight percept, as in Ryan 2011a. This second tack is briefly sketched here. First, a variation-capable constraint framework is assumed, such as maximum entropy harmonic grammar (maxent HG; e.g. Hayes & Wilson 2008) or noisy HG (e.g. Boersma & Pater 2015). A relevant constraint such as Weight-to-Stress, which penalizes unstressed heavy syllables (Prince 1983, Prince & Smolensky 2004 [1993]; cf. Stress-to-Weight and Peak-Prominence), can then be remolded in gradient fashion: 'For each unstressed syllable, increment the penalty by the duration of the p-center interval'.[20] This constraint is violated to a real number degree supplied by the perceptual-phonetic interface (on real-valued gradience in HG, see Flemming 2001, Katz 2010, and Ryan 2011a; see also Pater 2012 on the virtues of at least incrementally gradient constraints in HG). HG with mixed categorical and gradient constraints closely resembles the logistic regression models of §2–4. The weight of gradient Weight-to-Stress can amplify or damp the overall sensitivity to syllable weight, but cannot alter the contribution of the onset relative to that of the coda, which is intrinsic to the p-center interval.

To illustrate the mechanics of this framework, a tableau is sketched in Table 4 showing just two constraints (their weights estimated by regression) and two candidates (leaving aside additional candidates and prosodic structure for simplicity; cf. Kager 1999:151ff.). In addition to gradient Weight-to-Stress, categorical NonFinality (Prince & Smolensky 2004 [1993]), which (loosely speaking) disfavors final stress, is

[20] This interval can be taken to extend from the p-center either to the end of the rime or, adapting Steriade 2009 to the present proposal, to the p-center of the next syllable, if any (Steriade suggests that the weight domain spans the entire rime-to-rime interval). Furthermore, following Gordon (2002:60 et seq.), the appropriate phonetic function might not be duration alone but energy integrated over duration.

included and indexed to nouns as one possible implementation of the finality avoidance in that part of speech. For a fuller model, including competing gradient faithfulness constraints promoting durational stability, see Flemming 2001:35. The total penalty of each candidate is the sum of its weighted violations. Since this is a maxent tableau, these penalties are mapped onto probabilities by $e^{-penalty} / \sum e^{-penalty}$, where $e \approx 2.71828$ and the sum ranges over all candidates. This tableau therefore predicts 89% initial stress for the pseudoword *plizzoof*. The presence of gradient constraints does not necessarily preclude the coexistence of their categorical counterparts,[21] nor, indeed, of any of the previously developed categorical machinery for metrical phonology, including moras (Hyman 1985, Hayes 1989, Morén 2001).

/plɪzuf/	*p*	total penalty	NONFINALITY$_{noun}$ weight = 2.3	WEIGHT-TO-STRESS (gradient) weight = 1.8
a. ˈpl̥ɪ.zuf	0.889	0.675	0	0.375
b. pl̥ɪˈzuf	0.111	2.759	1	0.255

TABLE 4. Illustrative maxent tableau for two candidates for the pseudoword /plɪzuf/ exemplifying interacting real-valued and categorical violability.

For existing lexemes, whose stresses are normally fixed, it is necessary to override this kind of default probability distribution, either through higher-weighted markedness constraints (insofar as the fixed accents are predictable) and/or through prominential faithfulness (e.g. MAX-PROM: assign a violation for an underlying accent unrealized on the surface; Alderete 2001). For example, consider *móngoose* and *monsóon*. One possible analysis of the primary stress difference relies on underlying accent, for example, /mɑ́ngus/ and /mɑnsún/. Table 5 shows the operation of faithfulness in this case. Alternatively, one might analyze the difference as deriving from a less direct sort of memorization, namely, more abstract segmental representations, for example, /mɑngʊsɛ/ (cf. Chomsky & Halle 1968:45) and /mɑnsun/. Under this analysis, additional markedness constraints would do the work of MAX-PROM in Table 5. Nevertheless, nearly all analysts (including Chomsky and Halle (1968)) admit that listed exceptions exist in English, so highly weighted MAX-PROM is still needed. Similarly, even if MAX-PROM is operative for this pair, additional highly weighted markedness constraints are still needed to override certain other illicit stress possibilities in richness-of-the-base candidates, for example, a noun with accented final schwa (cf. Alderete 2001:21ff. on hybrid lexical/predictable stress systems). A more comprehensive discussion of English stress is beyond the scope of this article.

/mɑ́n.gus/	*p*	total penalty	MAX-PROM weight = 6.0	NONFINALITY$_{noun}$ weight = 2.3	WTS (gradient) weight = 1.8
a. ☞ ˈmɑ̃ŋ.gus	1	0.621	0	0	0.345
b. mɑ̃ŋˈgus	0	8.885	1	1	0.325

TABLE 5. Illustration of accentual faithfulness overriding the default probability distribution (though accent is not necessarily underlying in this case; see text). Secondary stress is omitted.

21 For example, Ryan 2011a:445 argues that weight mapping in the Homeric Greek hexameter requires both an impermeable binary distinction and intracategorial gradience within heavy syllables (though see Flemming 2001 for general arguments against gradient/categorical duplication).

The psychoacoustic orientation of this approach, in line with previous research in a similar vein (e.g. Gordon 2002, 2005), leaves open the possibility of additional phonetic contributions to the weight percept that are not specifically treated here, including the amplitude envelope (n. 20), tonal perturbation, and auditory recovery. Tonal effects include the frequent covariance of onset voicing and tone, with voicelessness favoring high tone (Yip 2002, Tang 2008, Kingston 2011), and high tone in turn favoring stress (de Lacy 2002). Microtonal perturbations might also affect gradient weight. Auditory recovery refers to the boost in the perception of the loudness of the rime following a low-intensity onset (Viemeister 1980, Delgutte 1982, Gordon 2005). Thus, the rime attack is more salient following a voiceless as opposed to voiced stop, a fact Gordon (2005) exploits to explain the onset voicing effect in languages such as Pirahã (Goedemans (1998:142, 148) expresses a similar insight without invoking recovery). Nevertheless, in addition to inducing greater recovery, voiceless onsets are longer, their p-centers occur earlier, and they engender (if anything) higher tone. These considerations are not mutually exclusive; it may be that both recovery and the p-center jointly effect the greater weight of voiceless onsets.

That said, however, auditory recovery cannot explain the full range of onset effects documented here. For example, recovery reaches ceiling at approximately 40 ms (Delgutte 1982:135), whereas events well outside of this window affect gradient weight. Consider the contrast between onset [b̥] and [sp] in Fig. 16. Given that the period of silence preceding the rime is over 80 ms in both cases, recovery predicts no weight difference. P-centers, by contrast, predict [sp] to be the heavier, evidently correctly (Fig. 6; see also Topintzi 2010:243 for a related criticism concerning onset geminates). The light weight of null onsets is likewise problematic for recovery, since this minimally energetic onset is expected to induce maximal, not minimal, salience of the rime attack. But p-centers make the right prediction about empty onsets (Fig. 17). Finally, while recovery is a low-level physiological phenomenon, onset weight effects are phonologized, emerging even in the absence of auditory exposure (e.g. in poetic composition and written wug tests). While one can tap along to syllables uttered silently to oneself, it is perhaps less clear that phonetic cognition also simulates auditory recovery absent the stimulation of the peripheral nervous system responsible for it (Delgutte 1982). In sum, auditory recovery is likely to be a factor in onset weight, but unlikely to be its sole explanation.

5.3. P-CENTERS: DISCUSSION. This section has put forth the hypothesis that the domain for syllable weight begins with the perceptual center, not the rime. Since no general theory of p-centers is yet agreed upon (see Villing 2010 for model comparisons), the underlying mechanics of p-center location, including the full range of relevant factors, is left as something of a black box here. Nevertheless, this article highlights an empirical aspect of p-centers that makes them superior to subsyllabic structure as candidates for delimiting the weight percept: p-centers are sensitive to onset properties, such that longer onsets tend to induce earlier p-centers, but only by a fraction of the onset's duration (§5.1).

Figure 18 depicts p-center and rime loci for *Brie* and *bee* with accompanying waveforms and spectrograms. As the figure makes explicit, both p-centers and rime onsets are somewhat ill-defined with respect to phonetic representations. For *Brie*, the most salient acoustic event is the release of *b*, not the relatively gradual transition from *r* into the rime. In this case, the p-center, judging from synchronization tasks (§5.1), is closer to the stop release. Whalen and colleagues (1989) find that even when participants are explicitly instructed to synchronize syllables so that the beginnings of the vowels are isochronous, their alignments are still more indicative of p-centers than of the actual vowel onsets. This apparent irrelevance of rime edges to syllable timing lends some a priori credence to a p-center account of weight in rhythmic weight systems, even putting aside the arguments from onset effects presented in this article.

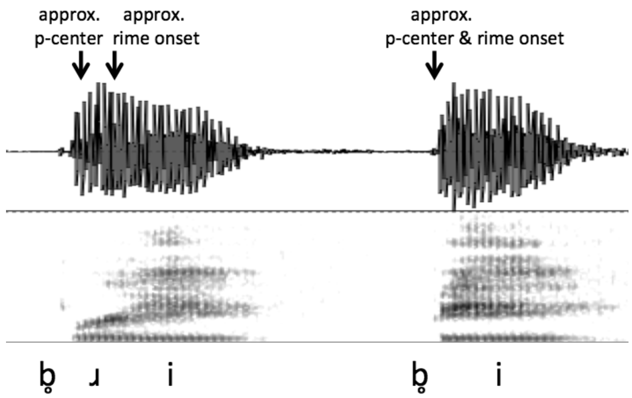

FIGURE 18. Approximate p-centers and rime onsets for *Brie* and *bee*.

In terms of typological implications, this approach predicts, for one, that onset effects should be universal in gradient weight systems of the type exemplified here. Second, it predicts rime structure to take precedence over onset structure in both gradient and categorical systems. As emphasized by Gordon (2005:600), this is the normal state of affairs for categorical weight. Pirahã (Everett & Everett 1984, Everett 1988) is a classic illustration of this principle. In its hierarchy of GV̌ < KV̌ < VV < GVV < KVV (where K is a voiceless consonant and G a voiced one), all short-voweled syllables are lighter than all long-voweled syllables. It is only within each rimal subset that onset properties come into play.

Nevertheless, as Gordon (2005) also observes, there exist apparent exceptions to rime primacy. In Arrernte, for instance, stress falls on a nonfinal peninitial syllable in just the case that the initial lacks an onset, ostensibly diagnosing a V < CV criterion. Stress skips an initial onsetless syllable even if it contains a coda; thus, VC < CV. Gordon (2005:625) provides a phonetic explanation for this apparent counterexample that is potentially compatible with the present account: controlling for stress, vowels in Arrernte are approximately twice as long in onsetful syllables as in onsetless ones. In fact, the absence vs. presence of an onset has a greater impact on the total energy of the rime than the absence vs. presence of a coda, making the former a superior binary discriminant even if weight is computed solely from the phonetics of the rime. This explanation carries over to the present account as long as the p-center shift conditioned on the onset is smaller than the rime-duration shift conditioned on the onset. This cannot be confirmed without Arrernte p-center data, but would not be unexpected given the small magnitudes of the p-center effects above. As a distinct approach, Arrernte stress has also been analyzed as quantity-insensitive. Downing (1998) and Goedemans (1998) use alignment constraints (McCarthy & Prince 1993) to render initial vowels of Arrernte polysyllables extraprosodic. For example, Goedemans (1998:168) posits ALIGN(foot, left, C, left), which requires feet to be consonant-initial. While these approaches, phonetic and phonological, might be invoked to cover Arrernte (and perhaps some related languages), they do not transfer to all of the categorical cases enumerated in §1.[22] Moreover, the present article does not depend on any categorical cases going through. Even if none were attested, the various onset effects in gradient weight systems would still require explanations that are beyond the scope of these particular arguments (e.g. as §2.1 makes clear, the onset effects in English stress cannot be attributed to interactions between onset properties and rime duration, as was possible for Arrernte).

With that caveat aside, however, the p-center account predicts that, insofar as categorical onset weight criteria are attested, the frequency of a particular criterion should correlate with its strength in gradient weight systems. This follows from Gordon's (2002) 'phonetic effectiveness' principle to the effect that an optimal binary criterion is one that maximally discriminates two groups of syllables along a perceptual continuum (as discussed further below, Gordon also invokes a simplicity bias). Consider two binary onset criteria, one based on presence vs. absence (Ø < C) and another on simplicity vs. complexity (C < CC). Among gradient systems, the cases examined here suggest that the former is generally the greater contrast.[23] Similarly, in the categorical typology (§1), Ø < C is the most frequent criterion, while there exist no clear cases of C < CC. A phonotactic confound might further inflate the incidence of the former criterion over the latter in the categorical typology: languages are more likely to permit null onsets than complex ones. In the World Phonotactics Database (Donohue et al. 2012), null onsets are permitted in 58% of languages, while complex onsets are permitted in 40% (N = 3,412). Complexity also seems to be more marked than nullity in that the former strongly implies the latter (with 88% accuracy), but not vice versa (41%).

A question then remains as to how the p-center effects discussed in this section might translate to a stronger Ø < C than C < CC contrast, given that both shifts appear to be roughly comparable for English in Fig. 17. I raise two possibilities here, both previously advocated in the syllable weight literature. First, the weight percept is plausibly sensitive not to linear increases in duration but to their PROPORTIONAL effect on the duration of the weight interval (following Weber's law; see Lunden 2006, 2011, and §4.1 above). For example, take a 100 ms syllable comprising only a vowel. Adding one coda consonant of 50 ms increases the syllable's duration by 50%; adding a second (also 50 ms) increases it by 33%. Thus, despite equal increments, the proportions attenuate. Lunden (ibid.) argues that weight must be understood in proportional rather than linear terms, consideration of which conceivably applies to the present case in which roughly incremental p-center shifts translate to differential contrasts.

A second (not mutually exclusive) possible explanation for the privilege of Ø < C over C < CC in both gradient and categorical weight systems concerns the arguably greater simplicity of presence vs. absence as opposed to degree criteria. Gordon (2002, 2006) maintains that weight criteria are grounded not only on phonetic effectiveness (see above), but also on formal simplicity. It is worth noting in this connection that Ø < C is also strongly overrepresented relative to C < CC in coda-based criteria (see also Ryan 2011a, e.g. p. 437, for cases of the former dominating the latter in gradient weight sys-

[22] Another phonological approach would be the stipulation, through constraints and their ranking, that onsets project moras in some languages but not others, independent of coda moraicity. As a referee points out, such an approach would straightforwardly cover apparent exceptions to rime primacy such as Arrernte, but otherwise fails to predict the near-universal asymmetry between onsets and codas in weight.

[23] For example, this is clear for English stress, as Fig. 1 shows. It is also true for Russian, though this is less obvious from Fig. 7 (cf. the discussion of 'floor effects' with respect to Fig. 4). Among Russian trisyllables, initial accent rates are 6.5%, 16.4%, and 22.1%, for Ø, C, and CC onsets, respectively (the jump from CC to CCC, for its part, is perhaps unexpectedly great, but still less in proportional terms than that from Ø to C).

tems). In the StressTyp database (Goedemans et al. 1996), by my count, eighty-three syllable weight criteria are sensitive to presence vs. absence of a coda, while only six are sensitive to coda complexity. Moreover, most if not all of those six cases rely not on coda complexity per se, but on the total number of timing slots in the rime (grouping $C_0\check{V}CC$ with C_0VVC as superheavy). While it is true that only about 25% of languages with codas permit complex codas (Donohue et al. 2012), this phonotactic skew is not in itself sufficient to motivate the degree to which nullity is favored over complexity in coda criteria. All of these observations appear to be paralleled by categorical onset criteria. The (arguable) absence of C < CC in categorical onset criteria might be an accidental gap, given the low frequencies of both onset-sensitive criteria and of languages permitting complex onsets. Especially if proportionality and/or simplicity are also in play, one would expect C < CC to be rare as an onset criterion even though its effects are robustly attested in gradient weight systems, for which optimal categorization is inapplicable.[24]

6. CONCLUSION. While onsets are rarely invoked by categorical syllable-weight criteria, their influence on weight emerges consistently in stress systems and quantitative meters exhibiting gradient variation. This influence is not predicted by rime-based theories of syllable weight. It is, however, predicted by the rhythmic theory of syllable weight proposed here, according to which the domain of weight begins not with the left edge of the rime but with the p-center, which approximates the left edge of the rime but is perturbed as a function of the onset. This proposal also predicts onsets to be subordinate to rimes as contributors to weight, given that rimes are parsed fully into the domain of weight while onsets are parsed into it only partially, if at all (§5). Thus, while onset effects are perhaps universal in gradient weight systems, they are typically masked by rime structure under small *n*-ary categorization.

REFERENCES

ALBRIGHT, ADAM, and BRUCE HAYES. 2006. Modeling productivity with the gradual learning algorithm: The problem of accidentally exceptionless generalizations. *Gradience in grammar: Generative perspectives*, ed. by Gisbert Fanselow, Caroline Féry, Ralf Vogel, and Matthias Schlesewsky, 185–204. Oxford: Oxford University Press.

ALDERETE, JOHN. 2001. *Morphologically governed accent in optimality theory*. London: Routledge.

ARCHANGELI, DIANA, and DOUGLAS PULLEYBLANK. 1994. *Grounded phonology*. Cambridge, MA: MIT Press.

ARCIULI, JOANNE; PADRAIC MONAGHAN; and NADA ŠEVA. 2010. Learning to assign lexical stress during reading aloud: Corpus, behavioural, and computational investigations. *Journal of Memory and Language* 63.180–96.

ARNOLD, EDWARD VERNON. 1905. *Vedic metre in its historical development*. Cambridge: Cambridge University Press.

BAAYEN, R. HARALD; RICHARD PIEPENBROCK; and LÉON GULIKERS. 1993. The CELEX lexical database [CD-ROM]. Philadelphia, PA: Linguistics Data Consortium, University of Pennsylvania.

BARBOSA, PLÍNIO; PABLO ARANTES; ALEXSANDRO R. MEIRELES; and JUSSARA M. VIEIRA. 2005. Abstractness in speech-metronome synchronisation: P-centres as cyclic attractors. *INTERSPEECH 2005*, 1441–44.

BECKER, MICHAEL. 2008. *Phonological trends in the lexicon: The role of constraints*. Amherst: University of Massachusetts Amherst dissertation.

[24] The strength of voicing contrasts relative to skeletal contrasts, for its part, is an area I leave to future research, in part because 'voicing' is realized rather differently across languages (§2.2) and the effects of these differences on p-centers is unclear.

BECKER, MICHAEL; NIHAN KETREZ; and ANDREW NEVINS. 2011. The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language* 87.84–125.

BECKER, MICHAEL; ANDREW NEVINS; and JONATHAN LEVINE. 2012. Asymmetries in generalizing alternations to and from initial syllables. *Language* 88.231–68.

BECKMAN, JILL N. 1998. *Positional faithfulness*. Amherst: University of Massachusetts Amherst dissertation.

BELTZUNG, JEAN-MARC. 2008. *Compensatory lengthening in phonological representations: Nature, constraints and typology*. Paris: University of Paris-3 (Sorbonne-Nouvelle) dissertation. Online: http://roa.rutgers.edu/, record 1026.

BERKO, JEAN. 1958. The child's learning of English morphology. *Word* 14.150–77.

BLEVINS, JULIETTE. 2004. *Evolutionary phonology*. Cambridge: Cambridge University Press.

BLUMENFELD, LEV. 2006. *Constraints on phonological interactions*. Stanford, CA: Stanford University dissertation.

BOERSMA, PAUL, and JOE PATER. 2015. Convergence properties of a gradual learning algorithm for harmonic grammar. *Harmonic grammar and harmonic serialism*, ed. by John J. McCarthy and Joe Pater. London: Equinox, to appear.

BOERSMA, PAUL, and DAVID WEENINK. 2011. Praat: Doing phonetics by computer. Version 5.2.23. Online: http://www.praat.org/, accessed May 1, 2011.

BOOIJ, GEERT. 1995. *The phonology of Dutch*. Oxford: Clarendon.

BROSELOW, ELLEN; SUSAN CHEN; and MARIE HUFFMAN. 1997. Syllable weight: Convergence of phonology and phonetics. *Phonology* 14.47–82.

BROWMAN, CATHERINE P., and LOUIS GOLDSTEIN. 1988. Some notes on syllable structure in articulatory phonology. *Phonetica* 45.140–55.

BURNAGE, GAVIN. 1990. *CELEX—A guide for users: English linguistic guide*. Nijmegen: University of Nijmegen, Centre for Lexical Information.

BURZIO, LUIGI. 1994. *Principles of English stress*. Cambridge: Cambridge University Press.

CARPENTER, ANGELA C. 2010. A naturalness bias in learning stress. *Phonology* 27.345–92.

CHEW, PETER A. 2003. *A computational phonology of Russian*. Oxford: Oxford University dissertation.

CHOMSKY, NOAM, and MORRIS HALLE. 1968. *The sound pattern of English*. Cambridge, MA: MIT Press.

COBUILD. 1987. *Collins COBUILD English language dictionary*. Glasgow: Collins.

COLÉ, PASCALE; ANNIE MAGNAN; and JONATHAN GRAINGER. 1999. Syllable-sized units in visual word recognition: Evidence from skilled and beginning readers of French. *Applied Psycholinguistics* 20.507–32.

CROSSWHITE, KATHERINE; JOHN ALDERETE; TIM BEASLEY; and VITA MARKMAN. 2003. Morphological effects on default stress in novel Russian words. *West Coast Conference on Formal Linguistics (WCCFL)* 22.151–64.

CUBBERLEY, PAUL V. 2002. *Russian: A linguistic introduction*. Cambridge: Cambridge University Press.

DAELEMANS, WALTER; STEVEN GILLIS; and GERT DURIEUX. 1994. Skousen's analogical modeling algorithm: A comparison with lazy learning. *Proceedings of the International Conference on New Methods in Language Processing*, 3–15.

DAELEMANS, WALTER, and ANTAL VAN DEN BOSCH. 2005. *Memory-based language processing*. Cambridge: Cambridge University Press.

DAELEMANS, WALTER; JAKUB ZAVREL; KO VAN DER SLOOT; and ANTAL VAN DEN BOSCH. 2010. *TiMBL: Tilburg memory-based learner reference guide, version 6.3*. Tilburg: Induction of Linguistic Knowledge Research Group, Tilburg Centre for Cognition and Communication.

DALAND, ROBERT; BRUCE HAYES; JAMES WHITE; MARC GARELLEK; ANDREA DAVIS; and INGRID NORRMANN. 2011. Explaining sonority projection effects. *Phonology* 28.197–234.

DAVIS, STUART. 1988. Syllable onsets as a factor in stress rules. *Phonology* 5.1–19.

DAYLEY, JON PHILIP. 1989. *Tümpisa (Panamint) Shoshone grammar*. (University of California publications in linguistics 115.) Berkeley: University of California Press.

DE LACY, PAUL. 2002. The interaction of tone and stress in optimality theory. *Phonology* 19.1–32.

DE LACY, PAUL. 2004. Markedness conflation in optimality theory. *Phonology* 21.145–99.

Delgutte, Bertrand. 1982. Some correlates of phonetic distinctions at the level of the auditory nerve. *The representation of speech in the peripheral auditory system*, ed. by Rolf Carlson and Bjorn Granström, 131–50. Amsterdam: Elsevier.

Deo, Ashwini. 2007. The metrical organization of Classical Sanskrit verse. *Journal of Linguistics* 43.63–114.

Donohue, Mark; Rebecca Hetherington; and James McElvenny. 2012. World Phonotactics Database. Canberra: Department of Linguistics, Australian National University. Online: http://phonotactics.anu.edu.au, accessed June 4, 2013.

Downing, Laura. 1998. On the prosodic misalignment of onsetless syllables. *Natural Language and Linguistic Theory* 16.1–52.

Eddington, David. 2000. Spanish stress assignment within the analogical modeling of language. *Language* 76.92–109.

Elbert, Samuel H. 1972. *Puluwat dictionary*. (Pacific linguistics C-24.) Canberra: Australian National University.

Everett, Daniel. 1988. On metrical constituent structure in Pirahã. *Natural Language and Linguistic Theory* 6.207–46.

Everett, Daniel, and Keren Everett. 1984. On the relevance of syllable onsets to stress placement. *Linguistic Inquiry* 15.705–11.

Fabb, Nigel, and Morris Halle. 2008. *Meter in poetry: With a chapter on Southern Romance meters by Carlos Piera*. Cambridge: Cambridge University Press.

Flemming, Edward. 2001. Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18.7–44.

Fougeron, Cécile, and Patricia Keating. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101.3728–40.

Francis, Winthrop Nelson, and Henry Kučera. 1982. *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.

Gabas, Nilson, Jr. 1999. *A grammar of Karo (Tupi)*. Santa Barbara: University of California, Santa Barbara dissertation.

Gahl, Susanne. 1996. Syllable onsets as a factor in stress rules: The case of Mathimathi revisited. *Phonology* 13.329–44.

Garde, Paul. 1976. *Histoire de l'accentuation slave*, vol. 1. Paris: Institut d'Études Slaves.

Goedemans, Rob W. N. 1998. *Weightless segments*. The Hague: Holland Academic Graphics.

Goedemans, Rob W. N.; Harry G. van der Hulst; and Ellis A. M. Visch. 1996. *Stress patterns of the world, part 1: Background*. (HIL publications 2.) The Hague: Holland Academic Graphics.

Goldman, Robert P. 1990. *The Ramayana of Valmiki: An epic of Ancient India: Balakanda*. Princeton, NJ: Princeton University Press.

Gordon, Matthew. 2002. A phonetically driven account of syllable weight. *Language* 78.51–80.

Gordon, Matthew. 2005. A perceptually-driven account of onset-sensitive stress. *Natural Language and Linguistic Theory* 23.595–653.

Gordon, Matthew. 2006. *Syllable weight: Phonetics, phonology, typology*. New York: Routledge.

Gouskova, Maria. 2010. The phonology of boundaries and secondary stress in Russian compounds. *The Linguistic Review* 27.387–448.

Gouskova, Maria, and Kevin Roon. 2009. Interface constraints and frequency in Russian compound stress. *Formal Approaches to Slavic Linguistics 17: The Yale meeting 2008*, ed. by Jodi Reich, Maria Babyonyshev, and Darya Kavitskaya, 49–63. Ann Arbor: Michigan Slavic Publications.

Gouskova, Maria, and Kevin Roon. 2013. Gradient clash, faithfulness, and sonority sequencing effects in Russian compound stress. *Laboratory Phonology* 4.2.383–434.

Guion, Susan G.; J. J. Clark; Tetsuo Harada; and Ratree P. Wayland. 2003. Factors affecting stress placement for English nonwords include syllabic structure, lexical class, and stress patterns of phonologically similar words. *Language and Speech* 46.403–27.

Gunkel, Dieter, and Kevin Ryan. 2011. Hiatus avoidance and metrification in the Rigveda. *Proceedings of the 22nd annual UCLA Indo-European Conference*, ed. by Stephanie W. Jamison, H. Craig Melchert, and Brent Vine, 53–68. Bremen: Hempen.

HAJEK, JOHN, and ROB GOEDEMANS. 2003. Word-initial geminates and stress in Pattani Malay. *The Linguistic Review* 20.79–94.

HALLE, MORRIS. 1973a. The accentuation of Russian words. *Language* 49.312–48.

HALLE, MORRIS. 1973b. Stress rules in English: A new version. *Linguistic Inquiry* 4.451–64.

HALLE, MORRIS. 1997. On stress and accent in Indo-European. *Language* 73.275–313.

HALLE, MORRIS, and MICHAEL KENSTOWICZ. 1991. The free element condition and cyclic versus noncyclic stress. *Linguistic Inquiry* 22.457–501.

HALLE, MORRIS, and SAMUEL JAY KEYSER. 1971. *English stress: Its form, its growth, and its role in verse*. New York: Harper and Row.

HALLE, MORRIS, and JEAN-ROGER VERGNAUD. 1980. Three dimensional phonology. *Journal of Linguistic Research* 1.83–105.

HALLE, MORRIS, and JEAN-ROGER VERGNAUD. 1987. Stress and the cycle. *Linguistic Inquiry* 18.45–84.

HAMMOND, MICHAEL. 1999. *The phonology of English: A prosodic optimality-theoretic approach*. Oxford: Oxford University Press.

HAYES, BRUCE. 1982. Extrametricality and English stress. *Linguistic Inquiry* 13.227–76.

HAYES, BRUCE. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20.253–306.

HAYES, BRUCE. 1995. *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

HAYES, BRUCE. 2012. How predictable is Italian word stress? Paper presented at National Chiao Tung University, Hsinchu City, Taiwan, May 11.

HAYES, BRUCE, and JAMES WHITE. 2013. Phonological naturalness and phonotactic learning. *Linguistic Inquiry* 44.45–75.

HAYES, BRUCE, and COLIN WILSON. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39.379–440.

HAYES, BRUCE; KIE ZURAW; PÉTER SIPTÁR; and ZSUZSA LONDE. 2009. Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85.822–63.

HOLM, STURE. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* 6.65–70.

HUBBARD, KATHLEEN. 1994. *Duration in moraic theory*. Berkeley: University of California, Berkeley dissertation.

HYMAN, LARRY. 1985. *A theory of phonological weight*. Dordrecht: Foris.

IVERSON, GREGORY. 1983. Voice alternations in Lac Simon Algonquin. *Journal of Linguistics* 19.161–64.

JAKOBSON, ROMAN. 1948. Russian conjugation. *Word* 4.155–67.

KAGER, RENÉ. 1989. *A metrical theory of stress and destressing in English and Dutch*. Dordrecht: Foris.

KAGER, RENÉ. 1999. *Optimality theory*. Cambridge: Cambridge University Press.

KATZ, JONAH. 2010. *Compression effects, perceptual asymmetries, and the grammar of timing*. Cambridge, MA: MIT dissertation.

KELLY, MICHAEL. 2004. Word onset patterns and lexical stress in English. *Journal of Memory and Language* 50.231–44.

KINGSTON, JOHN. 2011. Tonogenesis. *Blackwell companion to phonology*, vol. 4, ed. by Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume, and Keren Rice, 2305–33. Hoboken, NJ: Wiley-Blackwell.

KIPARSKY, PAUL. 1968. Metrics and morphophonemics in the Kalevala. *Studies presented to Roman Jakobson by his students*, ed. by Charles Gribble, 137–48. Cambridge, MA: Slavica.

KIPARSKY, PAUL. 2006. Amphichronic linguistics vs. evolutionary phonology. *Theoretical Linguistics* 32.217–36.

LEINO, PENTTI. 1994. The Kalevala metre and its development. *Songs beyond the Kalevala: Transformations of oral poetry*, ed. by Anna-Leena Siikala and Sinikka Vakimo, 56–74. Helsinki: Suomalaisen Kirjallisuuden Seura.

LIBERMAN, MARK. 1975. *The intonational system of English*. Cambridge, MA: MIT dissertation.

LIBERMAN, MARK, and ALAN PRINCE. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8.249–336.

LÖNNROT, ELIAS. 1849. *Kalevala taikka vanhoja Karjalan runoja Suomen kansan muinoisista ajoista*. Helsinki: Suomalaisen Kirjallisuuden Seura.

LUNDEN, ANYA. 2006. *Weight, final lengthening and stress: A phonetic and phonological case study of Norwegian*. Santa Cruz: University of California, Santa Cruz dissertation.

LUNDEN, ANYA. 2011. The weight of final syllables in English. *West Coast Conference on Formal Linguistics (WCCFL)* 28.152–59. Online: http://www.lingref.com/cpp/wccfl/28/index.html.

MACDONELL, ARTHUR ANTHONY. 1916. *A Vedic grammar for students*. Oxford: Oxford University Press.

MARCUS, STEPHEN MICHAEL. 1981. Acoustic determinants of perceptual center (P-center) location. *Perception & Psychophysics* 30.247–56.

MCCARTHY, JOHN J., and ALAN PRINCE. 1993. Generalized alignment. *Yearbook of Morphology* 1993.79–153.

MELVOLD, JANIS LEANNE. 1989. *Structure and stress in the phonology of Russian*. Cambridge, MA: MIT dissertation.

MORÉN, BRUCE T. 2001. *Distinctiveness, coercion and sonority: A unified theory of weight*. London: Routledge.

MORETON, ELLIOTT. 2008. Analytic bias and phonological typology. *Phonology* 25.83–127.

MORTON, JOHN; STEVE MARCUS; and CLIVE FRANKISH. 1976. Perceptual centers (P-centers). *Psychological Review* 83.405–8.

NANNI, DEBBIE. 1977. Stressing words in *-ative*. *Linguistic Inquiry* 8.752–63.

NEWCOMBE, ROBERT G. 1998. Two-sided confidence intervals for the single proportion: Comparison of seven methods. *Statistics in Medicine* 17.873–90.

NEWCOMBE, ROBERT G. 2000. Statistical applications in orthodontics, part II: Confidence intervals for proportions and their differences. *Journal of Orthodontics* 27.339–40.

OLDENBERG, HERMANN. 1888. *Die Hymnen des Ṛgveda I: Metrische und textgeschichtliche Prolegomena*. Berlin: Hertz. [Reprinted, Wiesbaden: Steiner, 1982.]

PATEL, ANIRUDDH D.; ANDERS LÖFQVIST; and WALTER NAITO. 1999. The acoustics and kinematics of regularly timed speech: A database and method for the study of the P-center problem. *Proceedings of the 14th International Congress of Phonetic Sciences (ICPhS)*, San Francisco, 405–8.

PATER, JOE. 2000. Non-uniformity in English secondary stress: The role of ranked and lexically specific constraints. *Phonology* 17.237–74.

PATER, JOE. 2012. Serial harmonic grammar and Berber syllabification. *Prosody matters: Essays in honor of Lisa O. Selkirk*, ed. by Toni Borowsky, Shigeto Kawahara, Takahito Shinya, and Mariko Sugahara, 43–72. London: Equinox.

PAYNE, DAVID, and FURNE RICH. 1988. Sensitivity to onset in Arabela stress. Pucallpa: Instituto Lingüístico de Verano, MS.

PETROVA, OLGA; ROSEMARY PLAPP; CATHERINE RINGEN; and SZILÁRD SZENTGYÖRGYI. 2006. Voice and aspiration: Evidence from Russian, Hungarian, German, Swedish, and Turkish. *The Linguistic Review* 23.1–35.

PITT, MARK; LAURA DILLEY; KEITH JOHNSON; SCOTT KIESLING; WILLIAM RAYMOND; ELIZABETH HUME; and ERIC FOSLER-LUSSIER. 2007. Buckeye Corpus of Conversational Speech. 2nd release. Columbus: The Ohio State University. Online: http://www.buckeyecorpus.osu.edu/.

PORT, ROBERT. 2007. The problem of speech patterns in time. *The Oxford handbook of psycholinguistics*, ed. by M. Gareth Gaskell, 503–14. Oxford: Oxford University Press.

PRINCE, ALAN. 1983. Relating to the grid. *Linguistic Inquiry* 14.19–100.

PRINCE, ALAN, and PAUL SMOLENSKY. 2004 [1993]. *Optimality theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell.

REVITHIADOU, ANTHI. 1999. *Headmost accent wins: Head dominance and ideal prosodic form in lexical accent systems*. The Hague: Holland Academic Graphics.

ROSE, YVAN. 2000. *Headedness and prosodic licensing in the L1 acquisition of phonology*. Montreal: McGill University dissertation.

RYAN, KEVIN M. 2011a. Gradient syllable weight and weight universals in quantitative metrics. *Phonology* 28.413–54.

RYAN, KEVIN M. 2011b. *Gradient weight in phonology*. Los Angeles: University of California, Los Angeles dissertation.

Sadeniemi, Matti. 1951. *Die Metrik des Kalevala-Verses*. Helsinki: Folklore Fellows Communications.

Seva, Nada; Padraic Monaghan; and Joanne Arciuli. 2009. Stressing what is important: Orthographic cues and lexical stress assignment. *Journal of Neurolinguistics* 22.237–49.

Sharoff, Serge. 2002. Meaning as use: Exploitation of aligned corpora for the contrastive study of lexical semantics. *Proceedings of Language Resources and Evaluation Conference (LREC02)*, 447–52. Online: http://www.lrec-conf.org/proceedings/lrec2002/.

Shelton, Michael. 2007. *An experimental approach to syllable weight and stress in Spanish*. State College: Pennsylvania State University dissertation.

Skousen, Royal. 1989. *Analogical modeling of language*. Dordrecht: Kluwer.

Skousen, Royal. 1992. *Analogy and structure*. Dordrecht: Kluwer.

Skousen, Royal. 2009. Expanding analogical modeling into a general theory of language. *Analogy in grammar: Form and acquisition*, ed. by James P. Blevins and Juliette Blevins, 164–84. Oxford: Oxford University Press.

Skousen, Royal; Deryle Lonsdale; and Dilworth S. Parkinson. 2002. *Analogical modeling: An exemplar-based approach to language*. Amsterdam: John Benjamins.

Sonderegger, Morgan, and Partha Niyogi. 2013. Variation and change in English noun/verb pair stress: Data, dynamical systems models, and their interaction. *Origins of sound patterns: Approaches to phonologization*, ed. by Alan C. L. Yu, 262–84. Oxford: Oxford University Press.

Soraghan, Christopher; Tomas Ward; Rudi C. Villing; and Joseph Timoney. 2005. Perceptual centre correlates in evoked potentials. *3rd European Medical and Biological Engineering Conference (EMBEC 2005)*.

Steriade, Donca. 2009. Units of representation for linguistic rhythm. Slides from the Linguistic Society of America Summer Institute, University of California, Berkeley, August 2009.

Tang, Katrina Elizabeth. 2008. *The phonology and phonetics of consonant-tone interaction*. Los Angeles: University of California, Los Angeles dissertation.

Tilsen, Sam. 2006. Rhythmic coordination in repetition disfluency: A harmonic timing effect. *UC-Berkeley Phonology Lab Annual Report*, 73–114.

Topintzi, Nina. 2010. *Onsets: Suprasegmental and prosodic behaviour*. Cambridge: Cambridge University Press.

van Nooten, Barend A., and Gary B. Holland (eds.) 1994. *Rig Veda: A metrically restored text with an introduction and notes*. (Harvard oriental series 50.) Cambridge, MA: Harvard University Press.

Viemeister, Neal. 1980. Adaptation of masking. *Psychological, physiological, and behavioural studies in hearing: Proceedings of the 5th International Symposium on Hearing*, ed. by G. van den Brink and F. A. Bilsen, 190–98. Delft: Delft University Press.

Villing, Rudi C. 2010. *Hearing the moment: Measures and models of the perceptual centre*. Maynooth: National University of Ireland, Maynooth dissertation.

Villing, Rudi C.; Tomas Ward; and Joseph Timoney. 2003. P-centre extraction from speech: The need for a more reliable measure. *Proceedings of the Irish Signals and Systems Conference (ISSC 2003)*, 136–41.

Vinogradov, Viktor Vladimirovich; Evgenija Samsonovna Istrina; and Stepan Grigorevich Barxudarov. 1960. *Grammatika russkogo iazyka*, vol. 1. 2nd edn. Moscow: Izdatel'stvo Akademii Nauk SSSR.

Westbury, John R., and Patricia Keating. 1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22.145–66.

Whalen, Douglas H.; André Maurice Cooper; and Carol Ann Fowler. 1989. P-center judgments are generally insensitive to the instructions given. *Phonetica* 46.197–203.

White, Laurence. 2002. *English speech timing: A domain and locus approach*. Edinburgh: University of Edinburgh dissertation.

Wilson, Edwin B. 1927. Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association* 22.209–12.

Wright, Matthew James. 2008. *The shape of an instant: Measuring and modeling perceptual attack time with probability density functions*. Stanford, CA: Stanford University dissertation.

Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.

Yoo, Seung-Nam. 1992. *The subsidiary stress in Russian compound words*. Urbana-Champaign: University of Illinois at Urbana-Champaign dissertation.

Yu, Alan C. L. 2011. On measuring phonetic precursor robustness: A response to Moreton. *Phonology* 28.491–518.

[kevinryan@fas.harvard.edu]

[Received 2 June 2012;
revision invited 11 February 2013;
revision received 26 March 2013;
accepted 20 May 2013]