Poznań Studies in Contemporary Linguistics 43(1), 2007, pp. 129–152

doi:10.2478/v10010-007-0007-2

PROMINENCE IN BEAT STRUCTURE

GEOFF SCHWARTZ
Adam Mickiewicz University, Poznań
geoff@ifa.amu.edu.pl

ABSTRACT

The sonority scale is generally invoked in order to provide an account of possible syllable or beat structures – the more sonorous a segment the more likely it is to serve as a nucleus. Auditory considerations, however, point to the primacy of onsets instead of nuclei as the most perceptually prominent segments in a syllable. Prominence Phonology (Schwartz, in preparation) considers the structure of a beat to be based on the interaction of two preference scales – one for onset prominence, and one for nuclear prominence. This paper presents the auditory properties of the two scales, and describes how the interactions between them may affect beat structure. These interactions can offer a plausible account of the "empty nucleus", and provide a formal vehicle in describing the application of Natural Phonological processes. The connection between beat structure prominence scales and the representation of segments in Prominence Phonology is also discussed, with implications for the interface between phonetics and phonology.

KEYWORDS: phonetics-phonology interface; Natural Phonology; element theory; auditory representations.

1. Introduction

One important aspect of Prominence Phonology (Schwartz, in preparation), a listener-oriented model of Natural Phonology (NP; cf. Stampe 1973; Donegan and Stampe 1979), is that it considers a variety of units that may be manipulated in phonological analyses, including feet, syllables, beats, and segments. However, of all of these units *beats* play perhaps the most pivotal role, providing the domain for the application of NP processes. The beat was promoted by Dziubalska-Kołaczyk (1995, 2002) in her Beats and Binding (B&B) framework. Beats are posited as the units that comprise a metrical foot – they are thus, as their name suggests, essentially rhythmic in nature. This is in opposition to syllables, which are seen as abstract constructs.[1] While syllables may be long or short, open or closed, beats are auditory constructs that have a rigid structure.

In B&B's formulation, a beat is represented by "a phoneme that is traditionally referred to as a syllabic nucleus" (Dziubalska-Kołaczyk 1995: 60). While there is a preference across languages for vowels as beats owing to their inherent sonority, consonants can occupy the beat position as well. In addition to beats, the B&B model posits what it refers to as *bindings*, which correspond to some extent to the syllabic notions of onset and coda. The motivation behind bindings is said to be found in the semiotic principle of "figure and ground" (e.g. Dressler 1985), which, when applied to phonology, emphasizes the need for sounds to contrast with neighboring sounds in order to be salient. Bindings are not posited as constructs in the present model. However, the notion of what bindings represent should fall out from the description of the auditory nature of beat onsets and nuclei.

The principle of figure and ground has a clear correlation with perception – objects are generally perceived in terms of their surroundings. Thus, when thinking about beats, syllables, or any other type of unit that a phonological framework may exploit, it is desirable to offer an explicit account of the auditory structure of the unit in question. To that end, this paper will consider certain aspects of the way the human auditory system processes acoustic stimuli. The goal is to outline the most basic auditory anatomy of a beat, allowing us to investigate the details of a listener-oriented version of Natural Phonology. In the model of NP presented here, every beat consists of an onset (ʔ) and a nucleus (@). These two primitives combine to form a single constituent. Codas have no status (cf. Lowenstamm 1996) in beat structure, but may emerge or evolve with language use.

The resemblance of the symbols adopted here to those employed in Element Theory (Harris and Lindsey 1995) is not accidental.[2] One of the most significant aspects of Element Theory is the notion that all elements are phonetically interpretable – there is no division between abstract underlying realizations and surface phonetic realizations. Thus, elements may represent a marriage of phonetics and phonology. If every phonological representation is pronounceable, then there is no place in a grammar for phonetic implementation – those phonetic phenomena that have been traditionally labeled "phonetic detail" must be representable in terms of categorical phonological primitives.

Space limitations prevent me from presenting a thorough case in favor of the role of phonetics in phonology. Convincing arguments can be found in Pierrehumbert (1999), who cites evidence from language specific differences in the realization of the same

---

[1] Beats and Binding Phonology indeed denies the syllable any status. Prominence Phonology posits syllables (along with segments) as "non speech units" that are not universal constructs but instead emerge with language use. Speech-based units (beats and feet) are the domain of NP "processes", while non-speech units may be manipulated by "rules".

[2] The symbol @ is used in e.g. Harris (1994) to refer to a neutral element, analogous to the carrier signal in Modulation Theory (Traunmuller 1994). In the present model the element @ fuses the neutral element with A.

sound, noting that "every thorough study that has looked for a difference between two languages in details of phonetic implementation has found one" (Pierrehumbert 1999: 115). The implication is that language-specific differences in phonetic detail are indeed systematic, and should be describable in phonological terms. In addition, psycholinguistic studies (see e.g. Lively et al 1994) have found that speakers' mental representations of speech sounds contain a lot of information that might be described by some as "phonetic detail". If phonology is meant to be concerned with mental representations of speech sounds, then much of this information should probably be considered phonological.

The present model will attempt to expand phonetic interpretability beyond the traditional feature-segment relations framed in terms of an element's independent realization. Essentially this will entail thinking of elements as "molecules", phonological substance that may be broken down into its component parts (perceptual cues). Just as the chemical behavior of water (a substance) cannot be understood without an understanding of the atomic structure of hydrogen and oxygen, the behavior of phonological elements cannot be understood without considering their auditory structure.

As mentioned earlier, this paper will be primarily concerned with two elements, an onset element ʔ and a nuclear element @. While most works have assumed that onsets and nuclei represent phonetically bare skeletal slots, the present model takes a different view. In accordance with phonetic interpretability, onset and nuclear elements must be specified in terms of some auditory realization. For each element, a set of 5 possible perceptual cues is posited. The realization of each element will be framed in terms of the number of cues present. For example, when 5 onset cues are projected by a segment, it is said to have an onset prominence of 5. Although most perceptual cues, as physical entities, are scalar in nature, for the sake of formal simplicity we shall adopt a one-to-one standard for elemental realization.

Many might see scalar primitives as incompatible with phonological analysis. Harris and Lindsey (1995) make the observation that "phonological oppositions, it is widely agreed, are inherently binary, rather than multi-valued or scalar" (Harris and Lindsey 1995: 37). However, Element Theory indeed unwittingly allows for scalar realization of its elements. The notion of "headedness" provides for three values in the realization of an element. An element is either absent, present as an operator, or present as a head. Although in theory headedness is determined by an element's position in the syntax of a phonological expression, in practice we are dealing with a 3-level scale of realization for the elements. Expanding this scale is necessary to describe cross-linguistic differences in the pronunciation of the "same" sound, and revive the prediction that every phonological representation is phonetically interpretable.

The focus of the present paper will be on the elements from which beats are constructed. We shall see that from an auditory point of view, the defining property of a beat is its onset. This is because onsets represent (pseudo-)discrete boundaries in the continuous speech stream. It is these boundaries that enable us to view the physical correlates of phonological categories – without onsets it is impossible to identify a nucleus.[3] Due to this primacy of onsets, the physical realization of what is generally called a "nucleus" can in some instances become secondary.

If it is the onset that defines a beat, there are also major implications for the notion of the "empty nucleus" proposed in the framework of Government Phonology (GP). This is a controversial proposal according to which each and every coda consonant represents the onset of empty nucleus. Although this proposal was born as a theory-internal characterization without any reference to phonetic properties, it is in fact quite provocative from an auditory-perceptual point of view. Naturalists have been reluctant to accept the empty categories of GP. However, the powerful predictions of the "empty nucleus" certainly warrant a second look to see if there is any compatibility with naturalist thinking. The auditory perspective offered here indicates that there is, although only in certain instances.

The rest of this paper will proceed as follows. Section 2 will discuss the auditory underpinnings of the perception of onsets. Section 3 will provide an explicit account of the realization of the onset element. Section 4 will provide an auditory perspective on silent nuclei. Section 5 will present an account of the realization of the nuclear element. Section 6 will discuss the nuclear qualities of sibilants. Section 7 will discuss how discuss the formation and representation of beats, as well as provide some final comments concerning the phonetics-phonology interface.

2. The perception of onsets

The perceptual motivation behind the primacy of onsets reflects response patterns in the auditory nerve.[4] The general pattern is that the onset of a stimulus receives an auditory boost, enhancing perceptual sensitivity. Thanks to this enhancement, listeners are readily able to identify the place of articulation of consonants in onset position. Consequently there is a tendency for phonological systems to allow more consonant contrasts in onset position than in coda position.

The auditory nerve processes acoustic phenomena in a phasic manner (Delgutte 1997). When an acoustic stimulus reaches the inner ear, auditory nerve fibres (ANFs) fire during a brief rapid adaptation phase of about 5 ms, producing what might be referred to as an onset boost or spike (Wright 2001). The rapid adaptation phase is fol-

---

[3] A *PSiCL* reviewer has charged that this is a case of "chicken and egg argumentation" and that the reverse statement could also be made, citing external place cues to onset consonants as an argument that a nucleus could define an onset. However, despite the important role of external cues, onsets may still be identified by their internal cues (e.g. the case of a stop produced "in isolation"). If a nucleus is to be defined, it must start somewhere. In other words, a nucleus is what follows an onset. However, an onset is defined not by the nucleus that follows it, but rather in terms of what precedes it.

[4] Similar patterns are observable throughout the auditory pathway to the level of the auditory cortex (Greenberg 1996).

lowed by a short-term adaptation phase, in which the heightened response level of ANFs slowly decays until a saturation point is reached. Although subject to variation, short-term adaptation lasts on the average about 50 ms (Wright 2004: 43). Taken together with the rapid adaptation, we may posit that auditory response is heightened over approximately the first 55 ms of an average beat or syllable. Such an interval corresponds roughly to the release burst of an unaspirated stop consonant followed by the formant transitions into the following vowel. In such cases both the stop release burst and the formant transitions will receive an auditory boost.

The actual point at which saturation is reached depends on the magnitude of the onset boost – the stronger the boost, the shorter the short-term adaptation, and saturation is reached sooner. The magnitude of the onset boost depends on rise time and amplitude of the stimulus, as well as the level of activity immediately preceding the onset (Wright 2004: 43) – less energy before the onset combined with a rapid rise time produce a higher magnitude onset boost. In other words, an onset is stronger if (1) it is preceded by silence and (2) it is marked by a rapid rise in amplitude.

The correspondence between onset prominence and the point of auditory saturation allows us to make certain predictions about rhythmic structure and phonotactic preferences. When onsets are low in prominence, saturation is delayed and more segmental material may align with a beat before the next onset boost. In such cases, codas may emerge and we can observe beats of different durations, features typical of stress-timed languages. When onsets are prominent and saturation occurs quickly, beats are more uniform in duration, intensifying the preference for CV structure.

If rapid and short-term adaptation can be associated with onsets, we may speculate that the saturation period which follows roughly corresponds to the nuclear position. During saturation, the firing rate of ANFs is lessened and auditory response is significantly reduced. If nuclei indeed align with this phase of auditory processing, then a nuclear segment, to reliably convey any linguistic information, must be sonorous enough to overcome saturation. As a result, there is a preference for vowels (or more sonorous segments) to fill the nuclear position. However, since the nucleus occupies a period of perceptual "numbness" during which auditory nerve fibres are recovering, it seems reasonable to assume that the crucial element in the structure of a beat is the onset boost. If such a boost is present, the stimulus can enjoy beat status even if it lacks a vocalized nucleus.

As an example let us consider the case of a stop consonant in final position. A final stop consonant consists of three portions are relatively distinct acoustically. The first portion occurs as the point of articulation is approached, producing formant transitions on the final portion of the vowel. The next portion is the closure, where acoustically we see a period of silence or near silence (with perhaps a small amount of periodic energy caused by voicing). Finally, the stop is released, producing a burst of aperiodic noise. These three distinct acoustic events are illustrated in Figure 1 below, illustrating a spectrogram of the Polish word *lek* 'medicine'. From an auditory perspective, the formant transitions coincide with the saturation period on the nucleus. Thus, if the stop is unreleased the only auditory information associated with it occurs on the vowel. During the closure period, the auditory system recovers from saturation, essentially "resetting" itself for the next onset. If the final stop is released, the release burst once again provides the auditory boost traditionally associated with onsets, regardless of whether or not the burst is followed by a vowel. If the burst is followed by silence, we are presented with a silent (although perhaps not "empty") nucleus. Returning to the final stop, of the three acoustic events, one (the VC formant transitions) is clearly attached to the preceding vowel, one (the release burst) is preceded by silence, and one (the closure) serves as a boundary between the two. This is clearly visible in the spectrogram in Figure 1 – we see two "bodies" of sound, and both of them contain cues for the /k/. It seems relatively safe to assume that the release burst of the /k/ will be amplified by the auditory system as an onset, and as such can be considered as a beat.

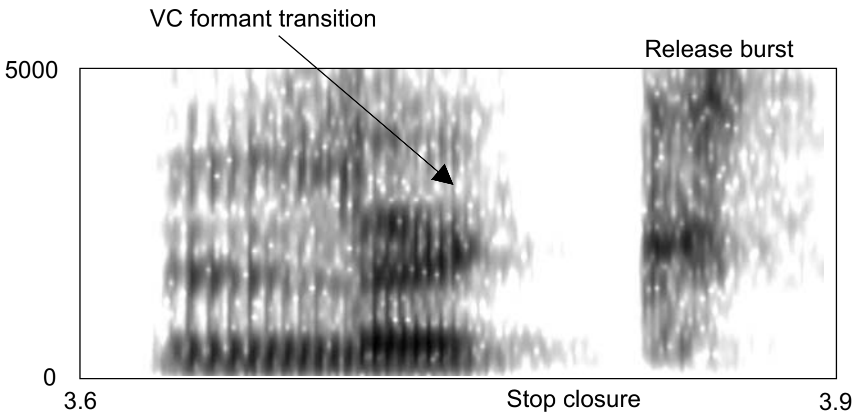

Figure 1. Spectrogram of the word *lek* 'medicine' produced by a speaker of Polish. Acoustic phenomena associated with the final /k/ are labelled. These include the VC transition (a merger of F2 and F3 for a velar), silence during stop closure, and the release burst.

The presence of the VC transition may provide some perspective on how codas may evolve according to a listener-based model of sound change (Ohala 1981; Blevins 2004). If the final stop is unreleased, the transitions of the /k/ are still present. Thus, listeners may associate cues to the final stop with the vocalic constituent. Such a scenario would be more probable when the VC transitions begin before auditory saturation, which is more likely when the preceding onset is less prominent. Listener/speakers may then generalize and introduce codas on a larger scale into a language. This type of development illustrates the view espoused in Prominence Phonology that beats are universal constructs containing only an onset and nucleus, while syllables are emergent units that may develop codas in certain conditions.

From an auditory perspective the apparent defining property of a beat is the presence of the onset boost, in which the initial portion of an acoustic stimulus is actually amplified by the auditory system. However, not all onsets are created equal – the magnitude of the onset boost varies with the type of articulation. Voiceless stops and affricates, in which auditory nerve recovery is made possible by a period of silence during closure, create very strong onset boosts (Delgutte and Kiang 1984). Voiced stops have shorter closure duration, allowing less auditory recovery time, and therefore show a less significant onset boost. The boost is even less significant for fricatives and nasals in coda position, since they do not provide the auditory nerve with the silent or near-silent period that allows for recovery from saturation. Thus, we might observe that the prominence of an onset is determined to some degree by what precedes it. One can compare onsets to the notion of *sforzando* in music. *Sforzando* is a sharp attack on the beginning of a note. In order for a note to be played *sforzando*, it must be separable from the note that precedes it. The same applies to onsets in speech.

While the onset boost is to some degree determined by what precedes the onset, inherent properties of the stimulus itself also play a role – some speech sounds are more sforzando-like than others. A "textbook case" of an onset boost is created by a sequence of glottal stop and a vowel, vowels with a glottal attack. This is because such a sequence has a rapid transition from silence to a sonorous vowel.

The preceding discussion points to two acoustic/auditory properties that are fundamental in the perceptual definition of an onset. The first of these is silence, and the second is a rapid rise in amplitude. While this would tend to point to a vowel with a glottal attack as the most prominent onset, additional onset cues tend to emerge with language input, since initial vowels with glottal attack provide for a somewhat impoverished inventory of possible lexical contrasts. An inventory of onset cues and a discussion of how they are realized segmentally is presented in the following section.

3. The realization of the onset element ?

The preceding discussion was not intended to provide a complete picture of the auditory processing of onsets – it was necessarily schematic in nature. It must be remembered, for example, that ANF's (as well as cells in the auditory cortex) have characteristic frequencies to which they respond. This fact complicates the task of modelling onsets in a phonological manner, since in principle auditory schemata would have to be created for a wide spectral range. At the same time, this fact strongly suggests that speech sounds are specified in the brain according to their physical properties, providing clear motivation for the incorporation of phonetics into phonology.

The auditory structure of onsets may be seen as a combination of four cues. In the Prominence Phonology model, a segment is said to project onset prominence settings based on the number of cues present when it is produced in onset position. Perceptual cues in this view are seen to be monovalent. Two of the cues for onset prominence were discussed in the previous section, and will be labelled Silence and RapidRise. While these two cues may be seen to be fundamental to the realization of onsets, two additional onset cues may be said to emerge from consonant articulations that inhabit onset positions. These cues are not inherent to onsets, but accompany the vast majority of Silence/RapidRise combinations. As a result they may be quickly incorporated into perceptual patterning. The first of these is aperiodic noise (Noise), and the second is formant movement (FM).

The prominence scale for ʔ is based on a one to one relationship between cues and prominence settings – each cue contributes one. The only caveat to this is that the RapidRise cue may in fact occur twice at the onset of a beat. This idea corresponds to Wright's (2001) two stages of onset enhancement, rapid adaptation and short-term adaptation. Rapid adaptation results in heightened sensitivity for roughly the first 5 ms of a stimulus, while short-term adaptation entails heightened but slowly declining sensitivity between rapid adaptation and the saturation. In the case of stop consonants, for example, both the release burst and the vowel onset may produce separate rises in amplitude. In such cases the burst rise receives a boost during the rapid adaptation period, while the vowel onset (and formant transitions) receive a rise during short-term adaptation. Thus, we shall posit RapidRise1 and RapidRise2 as separate cues to ʔ.

The case of fricatives as onsets requires additional discussion. Because the noise associated with fricatives tends to last longer than the adaptations associated with the auditory boost of onsets, one would expect saturation to set in at some point during the noise. However, it turns out that saturation is not merely a question of the duration of the stimulus, but whether or not there is periodicity in the signal. When the signal is periodic, auditory nerve activity tends to synchronize with that periodicity (up to about 5 kHz) in a process known as phase locking (Greenberg 1996; Delgutte 1997). Phase locking results in heightened auditory activity, leading to saturation (Wright 2004). Aperiodic frication noise on the other hand has been found to "distract the listener" (Bregman 1990, cited in Blevins and Garrett 2004: 128), delaying saturation and allowing for a boost at vowel onset even when the frication noise is long in duration.

Having posulated two separate rapid rise cues, we set the maximum realization of ʔ to 5, with a minimum of zero. Table 1 provides as summary of the cues and the articulations that produce them.

The preference scale of segment types in terms of onset prominence is presented in (1). It must be remembered that these settings are meant is preferences within the epistomology of Natural Lingustics – they are not meant to be absolute laws.

Table 1. Cues to the onset element ʔ

Cue	Description	Additional comments
Silence	Closure of stops and affricates	Voiced stops have shorter closure, voicing during closure may be perceptible.
RapidRise1 (RR1)	Release of stops and affricates	May not be present in lenis stops with short, weak bursts.
RapidRise2 (RR2)	Onset of vowel	May be present in fortified sonorants.
Noise	Aperiodic noise	Cue may not be present in voiced stops, especially /b/, that have minimal noise.
Formant Move ment (FM)	Coarticulatory effects of consonant with vowel	Always present except after glottal onsets

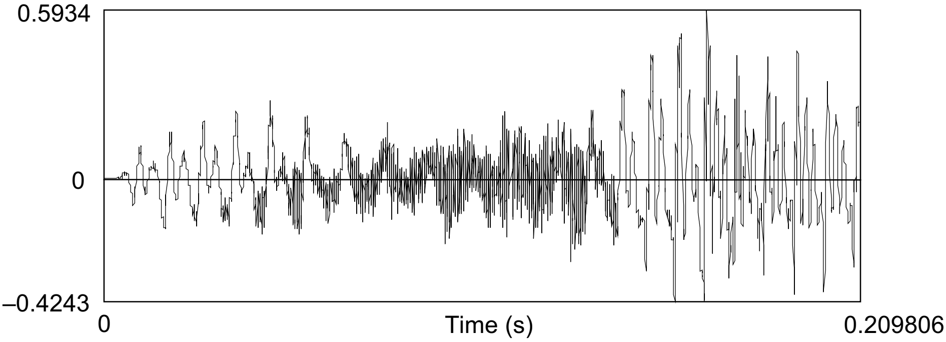

Figure 2. Waveform of initial fricative–vowel sequence in Polish *zielona* 'green'.

Voiceless stops and affricates feature all five of the cues in Table 1. Voiced stops are slightly less prominent as onsets. They may lack the silence cue when voicing during closure is strong. Alternatively, with very short and weak release bursts (especially in the case of /b/, see e.g. Maddieson 1997), they may lack the noise cue. Strident fricatives lack the silence and first Rapid Rise cue, and show an onset prominence of 3. The noise associated with non-strident fricatives may be quite weak, especially when voiced, so they may lack the noise cue. Voiced fricatives may lack RapidRise2. Voiceless fricatives feature long duration that provides time for a significant drop in amplitude before vowel onset, thus producing a rapid rise. Voiced fricatives often do not show a decrease in amplitude just before vowel onset. This can be seen in Figure 2, showing the fricative-vowel combination in the first syllable of Polish *zielona* 'green'.

Sonorant consonants may show a lot of variation in their realization of onset prominence. Most liquids and nasals feature FM and RR2. However, in many realizations of rhotics (e.g. uvulars), we may witness the additional cue of noise, producing an onset prominence of 3. On the other hand, liquids may be produced as approximants, in which case RapidRise2 may be absent. Approximants generally only contain FM. However, the glide /j/ may be produced with some frication and/or enough reduction in amplitude to produce a rapid rise in amplitude at vowel onset.

(1) Preference scale for onset prominence.

5	Voiceless (fortis) stops and affricates.
3–4	Voiced (lenis) stops and affricates (noise and/or silence may not be present).
2–3	Fricatives (Silence and RR1 are absent in all cases, RR2 or Noise may also be absent).
2	Liquids and nasals (RR2 and FM), vowels with glottal attack (Silence and RapidRise2), some non-strident fricatives.
1	Approximants (FM).

The scale presented above is reflected in a variety of phonological behavior, the most obvious being consonant lenition. Lenition will be formalized here as a decrease in the setting of onset prominence. Familiar examples include spirantization (a loss of RR1 and silence), intervocalic voicing (loss of RR2 in fricatives), and tapping (loss of silence).

3.1. Final comments on the primacy of onsets

Speech is a dynamic phenomenon. This fact presents a challenge to phonologists who subscribe to the notion that speech is relevant to phonological theory – the discrete units used in phonology are often very hard to isolate in the acoustic signal. However, certain phenomena – silence associated with stop closures, stop release bursts, vocalic onsets – *come close* to providing discrete boundaries in the signal. It therefore seems reasonable to assume that these acoustic events should be the most pivotal in delineating the discrete units of phonological theory. As we have seen in the previous discussion, such acoustic phenomena are intimately involved in the processes of speech perception – onsets correspond to a boost in auditory sensitivity, and therefore comprise the most important element of a beat. In accordance with the principle of figure and ground, the onset is foregrounded, seen against the backdrop of what precedes and what follows it. Due to the auditory primacy of the onset, the actual realization of the nucleus may become less crucial to the identity of a beat.

This notion exemplifies an area where there is a clear relationship between phonology and speech perception. Since most languages have a greater number of consonants than vowels, consonants tend to be far more important than vowels in the formation of lexical contrasts. Due to the fact that consonants are transients that are relatively low in perceptual salience, listeners often need to rely on the formant information in vowels in order to identify consonant place of articulation. Thus, the role of vowels in the transmission of consonants may serve to detract from the their own realization. This idea is borne out in phonostylistic processes – in careful speech vowel reductions are more likely to occur than consonantal lenitions (Malczak 2006). At the same time, in casual speech onsets tend to preserved, evident in forms like *p'tato*, *t'morrow, s'ppose* – forms that keep nuclei at the expense of onsets are not attested (**otato*, **omorrow*, etc.).

4. An auditory perspective on silent nuclei

In addition to an onset, every beat must contain is a nucleus. Most phonological theories involve a preference (or constraint) that nuclei should be vowels. This notion is of course closely related to the familiar sonority scale. Traditionally, sonority is cued by periodicity and formant structure in the acoustic signal. Vowels, of course, provide the most robust combination of these two acoustic phenomena, and thus occupy the highest position on the sonority scale. Sonorant consonants tend to have less robust formant structure (frequently with anti-formants or widened formant bandwidths, see Johnson 1997 for discussion) than vowels, and thus occupy a lower position. As such they are less likely to appear as syllabic nuclei. Obstruents have little or no periodicity, and generally lack formant structure, and as such occupy the lowest positions on the sonority scale. "Syllabic" obstruents are nearly unheard of across languages. From the perceptual point of view, nuclei occupy the period of auditory saturation – the portion of the beat from the end of short-term adaptation to the point of recovery. Thus, if a nucleus is to contain a segment capable of conveying linguistic meaning, that segment should preferably be sonorous enough to overcome auditory saturation.

If our auditory characterization of nuclei is to be complete, however, we need to consider two possible additional cues: duration and aperiodic noise produced by fricatives. As mentioned above, nuclei represent the period of auditory saturation, which lasts from the end of short-term adaptation, to the point of auditory nerve recovery. This period inevitably provides a window of time into which segmental material is placed. Different types of segment, of course, have different inherent durations. Consequently, some of them will provide a better fit into this window of time than others. This notion promotes two types of segments on an auditory scale of nuclear prominence. The first of these of course, are fricatives, since they are of inherently long duration. The following section will provide a detailed discussion on the auditory nature of fricatives and their place on a perceptual scale of nuclear prominence. The other type of "segmental material" that appears on an auditory scale of nuclear prominence, is, ironically, silence.

Figure 3 displays a spectrogram and waveform of the word *paczkę* 'package (Acc.)' in Polish. A traditional syllabification of this word would be *pacz.kę*, in which the affricate is a coda in the first syllable of a two-syllable word. Looking at the waveform, however, we immediately observe three distinct acoustic events separated by two periods of silence. The first is the stop-vowel combination. The second is the release burst and frication of the affricate. The third is the final stop-vowel combination. Temporally, the silent periods are quite prominent, and in each case would provide sufficient recovery time for the auditory system to prepare itself for the following onset. Notice as well that on the waveform, the affricate-silence combination is of comparable duration to the final "syllable" /ke/.

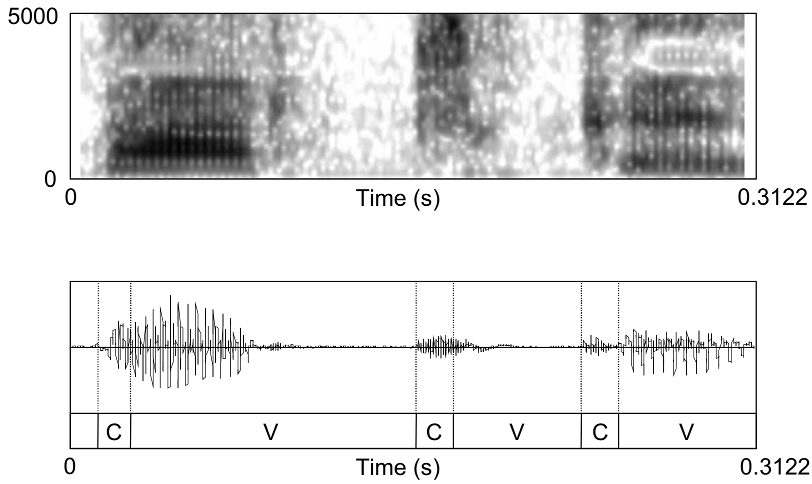

Figure 3. Spectrogram and waveform of the Polish word *paczkę*. In auditory terms, C corresponds with the period of boosted perceptual response associated with increased neural activity after stimulus onset, while V corresponds to the saturation period.

Thus, in a case such as this, the idea of a silent nucleus after the affricate is perceptually plausible – the silence in the closure period before the affricate release is sufficient for auditory recovery, and the release and frication of the affricate are prominent enough to qualify as an onset.

The above illustration raises the following question: when is a silent (empty) nucleus really a nucleus? From an auditory point of view, the answer should be clear from the figure – an empty nucleus is indeed a nucleus when it has a prominent onset. Stops and affricates in coda position certainly represent auditory onsets because they are preceded by closure. Sonorant consonants (nasals and liquids) in coda position, however, are not the best candidates for onset status since they lack closure.

Figure 4 below shows a waveform display of *z psem* 'with the dog' in Polish. The left and right-hand arrows mark the onset of the vowel and the nasal, respectively. Notice that there is no silent period preceding the /m/. Thus, this /m/ in coda position is unlikely to receive the auditory boost associated with onsets. As a result it is phonetically less plausible to posit silent nucleus in this case. Rather, it seems that the vowel and nasal are part of the same auditory unit.[5]

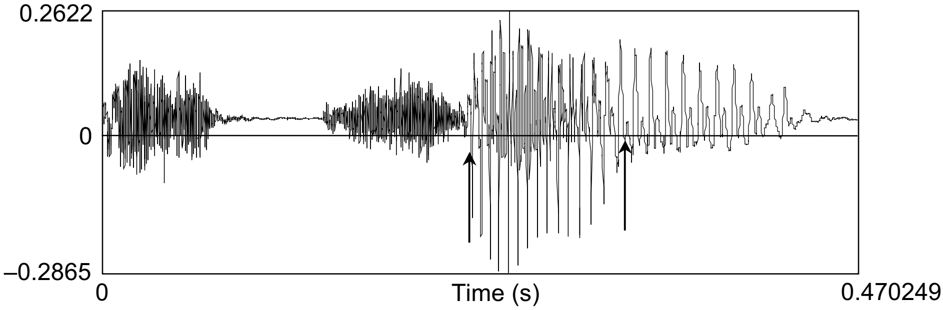

Figure 4. Waveform display of Polish *z psem* ' with the dog'. The left arrow denotes the onset of the vowel, while the right arrow denotes the beginning of the nasal consonant.

The prediction that falls out here is that post-vocalic consonants with low onset prominence should behave in a "coda"-like fashion, while those with high onset prominence should not. One area where this indeed occurs is in weight phenomena. Gordon (2004) observes that when languages make a weight distinction between CVR syllables (where R is a sonorant consonant) and CVC (with an obstruent coda), CVR syllables are always heavier. The present model provides a simple account of this distinction – CVC syllables are lighter because the coda consonant acts as an onset to a second beat. CVR syllables, on the other hand, generally contain three segments' worth of spectral material in a single beat. In this view, the primary ingredient in syllable weight is the alignment of beats with segmental material. We will return to the issue of codas in section dealing with guidelines for combining segments into beats, and the relationship of beats to syllables. Now, however, we must address the realization of the nuclear element @.

[5] It is possible to analyse the /m/ as its own beat, in which case it would span *both* onset and nucleus. I would speculate that the constituent structure of vowel–nasal combinations is a language-specific issue that is related to isochrony.

5. The realization of @

The element @ in Prominence Phonology represents a nucleus. In keeping with phonetic interpretability, every setting on the nuclear prominence scale must be specified perceptually in terms of cues. Each cue will be seen to contribute one to the prominence setting.[6] For the realization of @ we shall posit 5 cues: Loudness, Periodicity, Robust Formant Structure, Duration, and HighF1. These are presented in Table 2 below.

Table 2. Cues to the nuclear element @

Cue	Description/Comments
Periodicity	Vocal fold vibration. Periodic sounds are resistant to masking in the presence of background noise (Wright 2004).
Duration	Provides enhancement to the other nuclear cues.
Robust Formant Structure	Present in vowels and approximants, absent in nasals
Loudness	Present in sibilants, absent in chromatic vowels
HighF1	Energy in mid range harmonics corresponding to a high F1 contributes to subjective loudness (Crosswhite 2003). Present only in low vowels.

The preference scale for nuclear prominence is presented in (2). From the perceptual point of view, the more prominent the nucleus, the more likely it is to be perceived during the period of auditory saturation. Thus, nuclear prominence inevitably bears a resemblance to the traditional sonority hierarchy. However, unlike the sonority hierarchy, which is in essence a postulate, nuclear prominence is explicitly based on the presence of five monovalent cues.

Low vowels project all five of the cues for @. Non-low vowels lack only HighF1 with the exception of peripheral vowels may be missing the Loudness cue. Reduced vowels such as schwa lack HighF1 and the duration cue. Sibilants have a nuclear prominence of 2, with the loudness and duration. As with onset prominence, sonorant consonants are marked by a lot of variability. Liquids and approximants generally contain periodicity and formant structure. Nasals also contain periodicity. However, the formants in nasals have increased bandwidths that reduce their perceptual sali-

[6] For nuclear prominence it may be necessary to incorporate the implicational enhancement interactions of cues. For example, longer duration enhances loudness. However, for the sake of formal simplicity we shall maintain the monovalent status of cues.

ence, so robust formant structure is absent. On the other hand nasalization is a feature typically realized over long durations, so we can posit the presence of the duration cue in nasals.

(2) Preference scale for nuclear prominence

5	Low vowels.
3–4	Mid vowels and high vowels (HighF1 missing, Loudness may be missing).
3	Schwa and other extra short vowels (duration and high F1 missing).
2–3	Sonorant consonants (periodicity, robust formant structure (absent in nasals), duration (present in nasals).
2	Sibilants (loudness+duration present).
1	Other obstruents.

6. The nuclear qualities of sibilants

On the traditional sonority scale, fricatives occupy a low position, just above stop consonants. This classification is somewhat problematic. First of all, fricatives are segments with inherently long duration, a fact that cannot be overlooked if the rhythmic properties of speech are to be incorporated into our model. Secondly, the aperiodic noise produced by sibilant fricatives is quite intense. Acoustically, this noise produces a signal that is generally less intense than vowels. However, perceptually, because sibilants produce energy at frequencies to which the auditory system is quite sensitive, the aperiodic noise associated with sibilants may indeed be subjectively louder than many vowels.

The relationship between intensity and loudness can be seen in the Figures 5 and 6 below. Figure 5 displays the intensity contour and waveform of an /a/, a /ç/ and /ʂ̥/ as produced by a Polish speaker. As we can see the vowel has the greatest acoustic intensity of the three sounds (79.6 dB), followed by /ç/ (73.1 dB) and then /ʂ̥/ (71.1 dB). However, the relative subjective loudness of the three sounds is reversed as a result of their frequency characteristics. The human auditory system is most sensitive to loudness at frequencies between about 12 and 20 on the Bark scale (1700–5600 Hz). In the vowel, most of the energy is found at lower frequencies. However, the aperiodic noise produced by the sibilant fricatives falls exactly in this range.

This we can see in the cochlear spectrogram in the Figure 6. On the scale of subjective loudness (measured in sones), the following values were observed at the point of peak amplitude: /a/ – 35.4 sones, /ç/ – 44.6 sones, and /ʂ̥/ – 46.5 sones. Notice that in these examples, the loudest sound is indeed the lowest in amplitude.

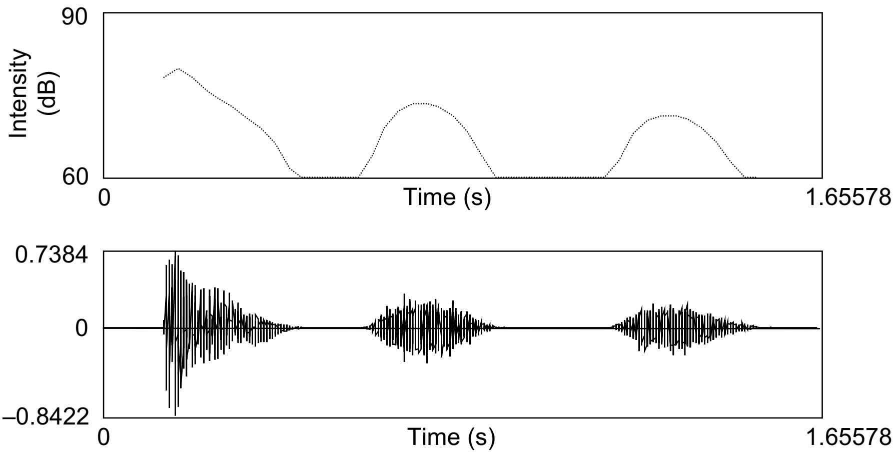

Figure 5. Waveform and intensity contour for three segments in Polish. The leftmost is /a/ (79.6dB), followed by /ɕ/ (73.1 dB), and finally /ʂ/ (71.1dB).

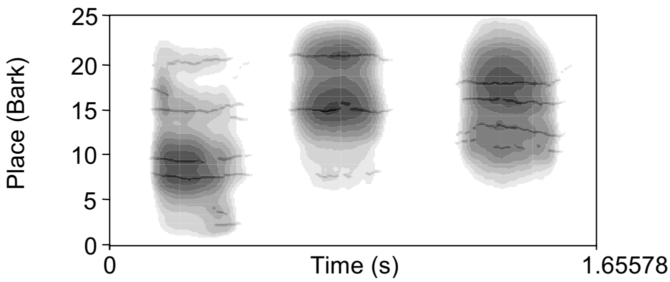

Figure 6. Cochlear spectrogram of the three Polish sounds in Figure 5. Subjective loudness measurements taken at peak amplitude were as follows: /a/ – 35.4 sones, /ɕ/ – 44.6 sones, and /ʂ/ – 46.5 sones.

While sibilant fricatives may be as loud or even louder than vowels, we cannot say for sure that they are more sonorous. This is because the perceptual cues transmitted by aperiodic noise are vulnerable to masking in noisy environments (Wright 2004). Thus, if we are to create a "perceptual sonority scale", determining the status of sibilant fricatives presents problems. These issues stem from the fact that certain acoustic/auditory properties (duration and noise intensity/loudness) associated with sibilants tend to promote their sonority, while the lack of periodicity detracts from their sonority.

The auditory properties of sibilants have implications for the representation of fricative-stop clusters. The position adopted here is similar to that of GP in that such clusters represent two beats. Granting sibilants nuclear status gives us a useful tool in tackling a number of problematic examples of phonological behavior. Sibilant clusters often behave differently than TR-type clusters.

First, SC clusters are notorious violators of the Sonority Sequencing Principle (SSP) (Selkirk 1984), which states that sonority should rise in onset clusters and fall in coda clusters, with a peak occurring on the nucleus. If the fricative were nuclear, the cluster then would contain two beats. In the first beat, the fricative fills both the onset (with a prominence of 2) AND the nucleus (with a prominence of 2) positions. This leaves the stop as a lone onset to the second beat. A quick look back at Figure 4 provides us with a useful perspective on the prospect that SC clusters represent two beats. On the waveform of Polish *z psem* [spsem] we see that the [sps] sequence neatly divides into two auditory events: [s] and [ps], with a clear break in between them – the [s] and [p] thus belong to different acoustic constituents. Consequently the stop can be seen to represent an independent onset, in which case the cluster would not constitute a violation of the SSP.[7]

Steriade (2006) compares SC clusters with TR clusters, noting that SC clusters tend to be more difficult to interrupt. For example, when English clusters are borrowed into Arabic, epenthesis breaks up the TR cluster, while it precedes the SC cluster. Thus the words *plastic* and *study* are realized in Arabic as [bilastik] and [istadi], respectively. A two-beat analysis of SC clusters explains this asymmetry. The stop closure already represents a discrete "interruption" of the cluster – there is little motivation to break it up further. Conversely, TR clusters present the listener with a continuous stream of sound with no discrete boundary. If the timing of the articulatory gestures is not coordinated, epenthesis may occur between the stop and the liquid, breaking up the cluster.

Polish adjectival morphology offers another example where nuclear status for sibilant fricatives would provide a solution to a difficult problem. Comparative adjectives in Polish are formed with the suffix *-szy* (e.g. *nowy–nowszy* 'new–newer'). If the stem ends in two consonants, however, an extra suffix *-ej-* precedes the *-szy* (e.g. *ładny–ładniejszy* 'pretty–prettier'). If, however, the stem ends in a /st/ cluster, then the *-szy* form is used (*prosty–prostszy* 'simple–simpler').[8] If the /s/ in the stem of *prostszy* is analysed as nuclear, for the purposes of morphology it could be equated with a vowel, explaining the origin of *prostszy*, a form that on the surface appears to be exceptional.

7. Representations, beat formation, and the phonetics-phonology interface

Natural Phonology has occasionally been criticized as being too phonetically oriented. Such a characterization fails to appreciate the NP view that "speech processing is cate-

---

[7] One might object to this analysis on the basis of an English word such as *stop*. If the /st/ cluster is two beats, then the /t/ can be seen as initial and should be subject to aspiration. However, in the two beat analysis of the /st/ cluster, the two beats are seen to comprise a foot, and the /t/ is not foot initial, and therefore is not subject to aspiration. Because of English phonotactic rules, aspiration is unnecessary after the /s/. Such a case illuminates the distinction between beats and syllables – with the consequence that in English the process of aspiration has graduated to become a rule.

[8] The form *czyściejszy* from *czysty* 'clean' is possible. However, *czystszy* is preferred.

gorical, or phonological, down to the level of the actual phonetic (pronounceable) representation" (Donegan 2002: 79). If the NP view is correct, the problem of the interface between phonetics and phonology should disappear, since phonological representations and processes could account for everything that many phonologists would attribute to phonetics. The implication is that phonology and phonetics are indeed a single domain.

In order to implement a merger of phonetics and phonology, two important tasks must be accomplished. First, it is necessary to build discrete mechanisms for representing gradient and dynamic phenomena. Prominence Phonology employs scalar representations for this purpose. The prominence scales proposed here, even if they turn out to need refinement, are still discrete. None of the scales proposed has more than five settings. While a five-setting scale may lack the computationally desirable characteristic of binarity, it should not be beyond the cognitive capabilities of human beings. After all, base ten is still the numerical system that we all use.

The other task is to create a system that relates cognitive units with the physical world. The model proposed here distinguishes between speech-based and non-speech-based units of analysis. Segments, while comprising probably the most basic cognitive unit in language, do not in themselves have any auditory reality – a segment out of context cannot exist in the physical world. If we observe the acoustic waveform of a single vowel "produced in isolation" we still see an auditory onset and formant structure that lasts a certain amount of time. Thus, what we observe in the physical world are beats. In the language learning process, however, segments are extracted from auditory input. As a result, we may assume that their cognitive representation includes some sort of auditory specification. Since the input to this representation is auditory, it must be assumed that all types of relevant auditory information is encoded. This includes not only information traditionally associated with the distinctive features of segments, but also higher-level information such as onsethood and nuclearity.

To relate cognitive segments with psychophysical beats, the representation of segments must therefore contain specifications for beat structure. In Prominence Phonology, the relations between segments, elements, and beats might be seen in similar terms as Turbidity (Goldrick 2000) that has been espoused in certain models of Optimality Theory. In Turbidity Theory, segments "project" structure (e.g. a feature, a mora), which may or may not be "pronounced". Structure that is present in the output, but lacking apparent motivation in the input is seen as "covert". In other words it is "projected" but not "pronounced".

In the present model, segments "project" prominence settings for elemental realization, while the realization of elemental prominence settings in the formation of beats may be thought of as "pronunciation". The terms employed by Goldrick seem indeed to be very apt in describing these relations in Prominence Phonology. Beats, because they are speech-based units, must be pronounced. Segments, as non-speech based units, are never pronounced. They do, however, project elemental structure that plays a role in beat formation. The following section will present the guidelines by which beat formation occurs.

7.1. From segments to beats and beyond

The formation of beats falls out from a small number of very simple principles in the Prominence Phonology model. Individual segments are assumed to project settings for onset and nuclear prominence according to the prominence scales given. When segments are combined in a string, beat structure is determined by two beat formation conditions. It is important to reemphasize the notion that although segments project beat structure prominence, they are not autonomous units in the representation of beats. Thus, a CV sequence is a single unit, a beat, which does not necessarily have a discrete boundary between the C and the V.

The beat formation rules outlined below represent only one of several forces that exert an influence on syllabic structure. Certainly there is a tendency across languages for beat structure and syllabic structure to correspond, reflected in the universality of CV syllable structures. However, the emergence of syllables may be subject to two other forces. The first of these is the influence of prosodic organization, which specifies the relative prominence of consecutive beats. The basic pattern, as observed by Bertinetto (1989: 108, cited in Dziubalska-Kołaczyk 1995: 57) is that stress-timed languages such as English show a tendency to reduce unaccented beats or syllables. Weaker beats are especially subject to onset lenition, in which case spectral material associated with lenited onsets may form codas and onset clusters. Analogy may also impact the emergence of syllable structure, since languages tend to group segments into consonant and vowel categories. These categories may be used instead of onset/nuclear prominence to produce phonotactic patterns that may not match the guidelines for beat formation. The discussion here is intended to concentrate on the role beat structure in the development of phonotactic patters – for further discussion of the role of prosodic organization, see Schwartz (in preparation).

The first rule and most basic of beat structure conditions is laid out in (3).

(3) An increase in onset prominence from one segment to the next constitutes a new beat.

This rule is the functional equivalent of the CVCV skeleton in GP. It makes the prediction that every coda consonant is in fact an onset. The effects of the condition are shown in a representation of the English word *bleak* in Table 3. The table presents each segment along with its projected onset prominence. Thus a word like this is analysed as two beats, and the final stop is considered an onset.[9]

[9] The vowel in this word would likely show formant transitions into the /k/. However, this is not to say that the /k/ belongs to both beats (cf. ambisyllabicity) in a formal sense. This is because segments have no status

Table 3. Onset and nuclear prominence for English *bleak*. According to the condition in (3), the /k/ constitutes an anonymous beat when it is released

Segment	Prominence of ʔ	Prominence of @
b	4	0
l	2	2
iː	0	3
k	5	0

The condition in (3) also provides for a two-beat analysis of sibilant-stop clusters, where sibilants represent autonomous beats. The /k/ in *bleak* projects a high onset prominence of 5. in accordance with the "rich get richer" principle we should expect that the "onsethood" of /k/ should be quite resistant to lenition. However, when "coda" consonants have a lower realization for ʔ, beat formation is somewhat more complicated.

From the perceptual perspective adopted here, final stops and affricates are the best candidates for "onset" status because their closure allows for auditory recovery that gives a boost to their release. While the condition in (3) stipulates that every increase in onset prominence represents a new beat, we need a way to describe the notion that final obstruents are better candidates for onsethood than final sonorants. This is handled with the view that natural processes are simply adjustments to the prominence scale. According to the "rich get richer" principle (or in this case poor get poorer), features that are present to only a small degree in segments may be especially subject to lenition. Such is the case with post-vocalic sonorant consonants, which generally have an onset prominence of two or less. When lenition occurs, it is not the segment that is lenited, but rather the segment's properties as an onset. In other words the segment's "onsethood" is lenited, but its spectral properties remain and must attach to the preceding nucleus. Thus, the emergence of codas can be seen as a resulting from onset lenition of a post-vocalic consonant.

In principle, a sequence of onset prominence increases may create consonantal beats that house only a single segment's worth of phonetic material. If such a beat is to be viable, it must be prominent enough to be perceptible. Since both onset and nuclear elements contribute (albeit in different ways) to the perceptual salience of a beat,[10] the prominence of a "single segment" beat may be seen in terms of the sum of its projected settings for ʔ and @. The condition is presented in (4):

in beat structure – beats only contain onsets and nuclei to which spectral properties such as formant transitions attach.

[10] See Gordon (2005) for a discussion of onset-sensitive weight.

(4) The sum of the onset and nuclear prominence of an autonomous beat must be at least 5.

This rule makes a number of predictions. The first is that coda obstruents signal the presence of an "empty nucleus", while sonorants generally do not. Another prediction that falls out from this rule is that sibilant-stop clusters receive a two beat analysis, while stop-liquid clusters do not. Conversely, initial liquid-obstruent clusters, which are extremely rare (cf. Scheer 1999), are possible only if the liquid is subject to either onset or nuclear fortition. A consequence of the condition in (4) is that as consecutive segments show onset increases, the onset prominence of all except the final segment is subject to lenition. This is clearly tied to phonetic context effects, because in those contexts that present consecutive increases in onset prominence, cues to ʔ may be absent.

When considering the notion in (4) it must be remembered that the onset and nuclear prominence scales presented earlier are posited as preferences rather than absolute rules. Thus, based on (4) we will not say categorically that a liquid (with ʔ–2, @–2 for total of 4) cannot serve as autonomous beat. However, if a liquid is to serve as an autonomous beat, either its onset or nuclear prominence must be fortified.

The effects of the condition in (4) can be seen in a representation of English *bark* in Table 4 below. In the beat formation of this word the onset prominence of /r/ is subject to lenition, with the consequence that /a/ and /r/ combine to form a single nucleus. The /k/ on the other hand, represents an autonomous beat.

Table 4. Onset and nuclear prominence representation for English *bark*; the /k/ represents an anonymous beat, while the /r/ is subject to onset lenition

Segments	ʔ	@
b	4	0
a	0	5
r	2>0	2
k	5	0

7.2. A note on process application and prosodic organization

Different process applications fall out from the correlation between prosodic organization and the relative prominence of the beats in a language. Stress languages such as English feature significant reduction of unstressed syllables. Such reduction is minimized in non-stress languages such as French. Thus, the role of prosodic organization in beat structure can be observed in a comparison of phonetically "similar" words in different languages. Consider a sequence /pul/ that in English produces the word *pool* but in French produces *poule* 'hen'.[11] Each of these words would have the same *projected* beat structure specification, however, the /l/ in the French word behaves more like an onset. Thus, in French the /l/ beat is more likely to be preserved. Since onsets are seen as the defining property of a beat, preserving a beat entails fortifying its onset properties. Schwa epenthesis makes the /l/ more prominent as an onset by adding a Rapid Rise cue and formant transitions. In English the final /l/ loses its onset properties (and may even be vocalized), reflecting its status as a weak, unstressed beat.

8. Final remarks

Although the sonority hierarchy is very commonly invoked to explain a range of phonological phenomena, it is often criticized as well. One of the main criticisms is that there is not an explicit way of defining sonority physically, leading sonority proponents to resort to circular reasoning in defending their proposals. The present model attempts to remedy this problem by providing a more concrete auditory perspective on the phenomena that sonority is invoked to account for. From this perspective it is more fruitful to look at onsets rather than nuclei, as nuclei often align with auditory saturation while onsets occur during periods of heightened perceptual sensitivity. Onset prominence may thus be viewed as "anti-sonority", an explicit auditory scale that provides perspective on the development of phonotactic patterns.

This paper is subject to space limitations, and is admittedly somewhat impoverished when it comes to empirical support for the proposals. However, clear guidelines have been established for further empirical study, enabling the inquisitive reader to attack both new and familiar data from the perspective adopted here.

REFERNCES

Bertinetto, P.M. 1989. "Reflections on the dichotomy 'stress v syllable timing'". *Revue de phonetique appliqué* 91/93. 99–130.

Blevins, J. and A. Garrett. 2004. "The evolution of metathesis". In Hayes, B., R. Kirchner, and D. Steriade (eds.). *Phonetically based phonology*. Cambridge: Cambridge University Press. 117–156.

Blevins, J. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.

Bregman, A.S. 1990. *Auditory scene analysis*. Cambridge, MA: MIT Press.

[11] For illustrative purposes we shall ignore the "long" specification of the English vowel. The length specification would actually enhance the argument here, as it reflects the tendency in English for more prominent accented syllables and highly reduced unaccented ones.

Crosswhite, K. 2003. "Spectral tilt as a cue to word stress in Macedonian, Polish, and Bulgarian". *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona*. 767–770.
Delgutte, B. 1997. "Auditory neural processing of speech". In Hardcastle, W. J. and J. Laver (eds), *The handbook of phonetic sciences*. Oxford: Blackwell Publishers. 509–538.
Delgutte, B. and N. Kiang. 1984. "Speech coding in the auditory nerve IV: Sounds with consonant-like dynamic characteristics". *Journal of the Acoustical Society of America* 75(3). 897–907.
Donegan, P.J. 1979. On the Natural Phonology of vowels. [PhD Dissertation, Ohio State University.] *Ohio State University Working Papers in Linguistics* 23. [Republished 1985. New York: Garland Publishing Company.]
Donegan, P. 2002. "Phonological processes and phonetic rules". In Dziubalska-Kołaczyk, K. and J. Weckwerth (eds.), *Future challenges for Natural Linguistics*. Munich: Lincom. 57–81.
Dressler, W.U. 1985. *Morphonology*. Ann Arbor: Karoma.
Dziubalska-Kołaczyk, K. 1995. *Phonology without the syllable*. Poznań: Motivex.
Dziubalska-Kołaczyk, K. 2002. *Beats and Binding Phonology*. Frankfurt am Main: Peter Lang.
Goldrick, M. 2000. "Turbid output representations and the unity of opacity". *Proceedings of NELS* 30. 231–245. [Also available as ROA 368-1299.]
Gordon, M. 2004. "Syllable weight". In Hayes, B., R. Kirchner and D. Steriade (eds.), *Phonetically based phonology*. Cambridge: Cambridge University Press. 277–312.
Gordon, M. 2005. "A perceptually-driven account of onset-sensitive stress". *Natural Language and Linguistic Theory* 23. 595–653.
Greenberg, S. 1996. "Auditory processing of speech". In Lass, N. (ed.), *Principles of experimental phonetics*. St. Louis: Moseby-Year. 362–408.
Harris, J. 1994. *English sound structure*. Oxford: Blackwell.
Harris, J. and G. Lindsey. 1995. "The elements of phonological representation". In Durand, J. and F. Katamba (eds.), *Frontiers of phonology: Atoms, structures, derivations*. Harlow, Essex: Longman. 34–79.
Johnson, K. 1997. *Acoustic and auditory phonetics*. Cambridge, MA: Blackwell.
Lehiste, I. and G.E. Peterson. 1959. "Vowel amplitude and phonemic stress in American English". *Journal of the Acoustical Society of America* 31. 428–435.
Lively, S.E., D.B. Pisoni and S.D. Goldinger. 1994. "Spoken word recognition: Research and theory". In Gernsbacher, M. (ed.), *Handbook of psycholinguistics*. New York: Academic Press. 265–301.
Lowenstamm, J. 1996. "CV as the only syllable type". In Durand J. and B. Laks (eds.), *Current trends in phonology: Models and methods*. Salford, Manchester: ESRI. 419–441.
Maddieson, I. 1997. "Phonetic universals". In Hardcastle, W. and J. Laver (eds.), *The handbook of phonetic sciences*. Cambridge: Blackwell. 619–640.
Malczak, K. 2006. The implicational nature of phonostylistic processes. [Unpublished MA thesis, Adam Mickiewicz University, Poznań.]
Ohala, J. 1981. "The listener as a source of sound change". In Masek, C. S., R. A. Hendrik and M. F. Miller (eds.), *Papers from the parasession on language and behavior*. Chicago: Chicago Linguistic Society. 178–203.
Pierrehumbert, J. 1999. "What people know about sounds of language". *Studies in the Linguistic Sciences* 29(2). 111–120.
Schwartz, G. In preparation. Phonology for the listener and language learner.
Selkirk, E. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

Stampe, D. 1973. A dissertation on Natural Phonology. [Unpublished PhD dissertation, University ofChicago.] [Republished with annotations 1979. New York: Garland Publishing Company.]

Steriade, D. 2006. Contour correspondence: evidence from cluster interruption. Talk given at the Third Old World Conference in Phonology. Budapest.

Traunmüller, H. 1994. "Conventional, biological, and environmental factors in speech communication: A modulation theory". *Phonetica* 51. 170–183.

Wright, R. 2001. "Perceptual cues in contrast maintenance". In Hume, E. and K. Johnson (eds.), *The role of speech perception in phonology*. San Diego: Academic Press. 251–277.

Wright, R. 2004. "Perceptual cue robustness and phonotactic constraints". In Hayes, B., R. Kirchner and D. Steriade (eds), *Phonetically basedphonology*. Cambridge: Cambridge University Press. 34–57.