SPECIAL COLLECTION: HEADEDNESS IN PHONOLOGY

Formalizing modulation and the emergence of phonological heads

Geoffrey Schwartz

Faculty of English, Adam Mickiewicz University, Al. Niepodległości 4, 61-874, Poznań, PL
geoff@wa.amu.edu.pl

This paper provides an account of how certain instances of "headedness" in segmental phonology may be derived within the Onset Prominence (OP) representational framework. It is shown that headedness is not a primitive property of OP representation, but rather emerges directly from the phonetic anatomy of the OP representational primitives, envisioned in terms of Traunmüller's Modulation Theory. The phonological status of voicing, including the relationship between nasals and voiced stops has been ascribed to headedness. Here it is shown to fall out from the Modulation perspective on laryngeal phonology. With regard to vowel quality, it is shown that apparent headedness effects derive from asymmetries in the modulatory properties of formant convergences as opposed to individual formants. Empirical implications of this perspective are reflected in vowel harmony patterns, by which rounding is typically less likely to be harmonic than palatality or tongue root advancement.

1 Introduction

In phonological representation, the term "head" is commonly invoked to refer to an entity that is dominant over other elements in a given structure. The label is most frequently encountered in the study of metrical or suprasegmental structure (e.g. Hayes 1995), where it is commonly used as a label for the most prominent foot in a word, or the most prominent syllable in a foot. Thus, the standard view is that stress is assigned to the "head" of a domain. For example, the English word *Massachusetts* is comprised of two feet, each of which is made of two syllables (*mæ.sə*)($\widehat{t\int}$u:.*səts*). These feet are trochaic, with the leftmost syllable as the head. On the word level, the head is the rightmost foot, so the primary word stress is located on the first syllable of the second foot (the third syllable in the word).

Heads are also invoked in descriptions of syllable-internal structure, in which the most sonorous vocalic element is typically assigned the label "nucleus" or "peak" that is assumed to be dominant with regard to consonants. For example, Smith's (2002) formulation of the ONSET constraint is encoded as a requirement that the head segment must not be the first segment in the syllable. This formulation is given in (1).

(1)   Formulation of ONSET (after Smith 2002)
        For all syllables σ, *a* ≠ *b* where *a* is the leftmost segment dominated by σ, *b* is the head segment of σ

Smith's approach is interesting in that it does not require the head of the syllable to be a vowel – it merely states that a syllable cannot be head-initial.[1] Nevertheless, it takes as a given that a syllable must have a head.

The final area in which heads appear in phonology is segmental representation. Some theorists have proposed that in the representation of individual phones there may be a single dominant feature. Early formal developments of this idea are found in Particle Phonology (Schane 1984) and Dependency Phonology (Anderson & Ewen 1987), and later the concept is used widely in Element Theory (ET; Harris & Lindsey 1995). In these frameworks vowel quality is represented in terms of combinations of monovalent primes for backness/rounding {U}, palatality {I}, and openness or sonority {A}. These components have been particularly insightful for characterizing the representation of vowel contrasts. In particular, mid vowels may be seen as being composed of a combination of the sonority prime with either the backness or frontness prime. When mid vowels contrast (e-ɛ; o-ɔ), headedness is used to encode the distinction between vowels that are in close proximity in acoustic vowel space.

The phonetic shape of phonological heads is typically thought of in terms of acoustic prominence. In the domains of metrics and syllable structure, this is observable in such measures as amplitude, pitch, duration, and spectral balance. In the case of vowel quality, heads are described in terms of formant frequencies, with peripheral values denoting headedness. Such metrics are for the most part reliable in determining what is a head and what is not a head. However, it is less clear whether and how these phonetic and functional considerations should be captured in a formal theory of phonological representation. An insightful perspective in this regard is provided by Modulation Theory (MT; Traunmüller 1994), according to which the linguistic (rather than extra-linguistic) elements in speech are seen as modulations on a carrier signal. From this perspective, phonological primes are determined not so much on the basis of their own specific articulatory or acoustic properties as on the basis of their modulatory effects on the carrier. Based on what we know about speech perception, it is reasonable to suggest that carrier modulations have a categorical aspect to them that lends itself to phonological interpretation. Thus, by considering the question of headedness from the perspective of Modulation Theory, we may explain the emergence of headedness effects in the domain of segmental representation.

To achieve this goal, it is necessary to provide an explicit formal expression of the carrier and its modulations. Traunmüller's original proposal envisions the carrier signal as a vocoid with evenly spaced formant frequencies characteristic of the vowel schwa. In what follows, I shall propose a more detailed perspective on the auditory properties of carrier modulation. In the area of vowel quality, the auditory structure of the elements {I} and {U} emerges from a set of privative spectral modulations. Vowel headedness effects reflect the presence of a greater number of salient modulations, encoding the relationship between formant convergences and the frequencies of single formants. In the case of consonants, salient modulations associated with manner of articulation are observable in the amplitude envelope as acoustic landmarks (Stevens 2002). In particular, stop closures and aperiodic noise associated with obstruents, especially in the absence of periodicity, produce robust modulations on the carrier. Finally, if we take Modulation Theory seriously and posit that voicing is part of the carrier, we must assume that voiceless consonants are characterized by the most salient modulations, regardless of the "type" of laryngeal system in terms of voicing or aspiration. Thus, voiceless consonants are always phonologically specified with respect to voiced consonants since they produce more robust perceptual effects on the carrier.

---

[1] Smith gives a number of arguments in favor of this formulation based on the ambiguous behavior of glides, and the desirability of eliminating the need for prosodic constraints to make direct reference to segmental features or sonority.

This paper will outline the perspective on these issues afforded by the Onset Prominence framework (OP; Schwartz 2010 et seq.), which provides representational materials to incorporate each of these aspects of Modulation Theory. Along the way, we shall consider how empirical effects that have been attributed to headedness may arise. One such case looks at the relationship between voicing and nasality. Another involves asymmetries in vowel harmony patterns with regard to the spreading of palatal vs. labial features, and the interaction between rounding and vowel height. In both instances, it will be shown that there is no need for a formal device to encode headedness effects, which emerge directly from the auditory structure of phonological primitives in the OP environment.

2 Amplitude modulation as prosodic structure

This section will provide a brief introduction to the Onset Prominence representational environment (Schwartz 2010 et seq.). For more thorough presentations, see Schwartz (2013, 2016a). For a preliminary presentation of OP laryngeal phonology, see Schwartz (2016b)

OP builds on earlier work exploring the hypothesis that manner of articulation is a structural property (Steriade 1993; Golston & van der Hulst 1999; Pöchtrager 2006). Individual segmental representations are extracted from a hierarchical structure derived from the phonetic events associated with a CV sequence in which the consonant is a stop. The top node (Closure) is derived from stop closure, the Noise node from aperiodic noise associated with frication and release bursts, and the Vocalic Onset (VO) node captures periodicity with formant structure associated with CV transitions as well as sonorant consonants. The Vocalic Target (VT) node houses (more or less) stable formant frequencies that define vowel quality. The hierarchy is presented in (2).

(2) The Onset Prominence representational hierarchy
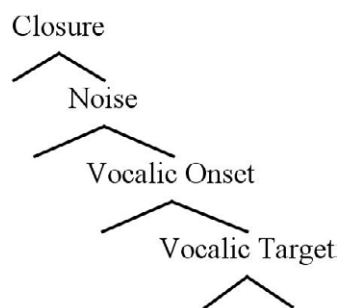
An important aspect of the structure in (2) is that linear order falls out directly from the sequence of phonetic events encoded in the CV (cf. Golston & van der Hulst 1999). To visualize this, consider Figure 1, in which the labels for the structural nodes in the OP are projected onto a waveform and spectrogram display of a stop-vowel sequence. In a stop-vowel sequence, Closure precedes Noise, which precedes the onset of the vowel (VO), which in turn precedes the vocalic target (VT). When the consonant is not a stop, individual phonetic events are missing (Closure in the case of fricatives, Noise in the case of nasals, Closure and Noise in the case of approximants), but the basic sequence is universal. Closure, if present, is always first. VT is always last. VO precedes only VT. Noise is first if Closure is absent, etc.

In waveform and spectrogram displays such as those in Figure 1, time is represented in what looks to be a linear fashion, from left to right. It therefore is fair to ask why the proposed CV primitive is a hierarchical structure rather than a simple linear string of prosodic positions. The answer to this question stems from the causal relationships inherent between the articulation of speech sounds and their acoustic consequences. The articulation of a stop-vowel sequence, which we assume to be the only universal prosodic unit,
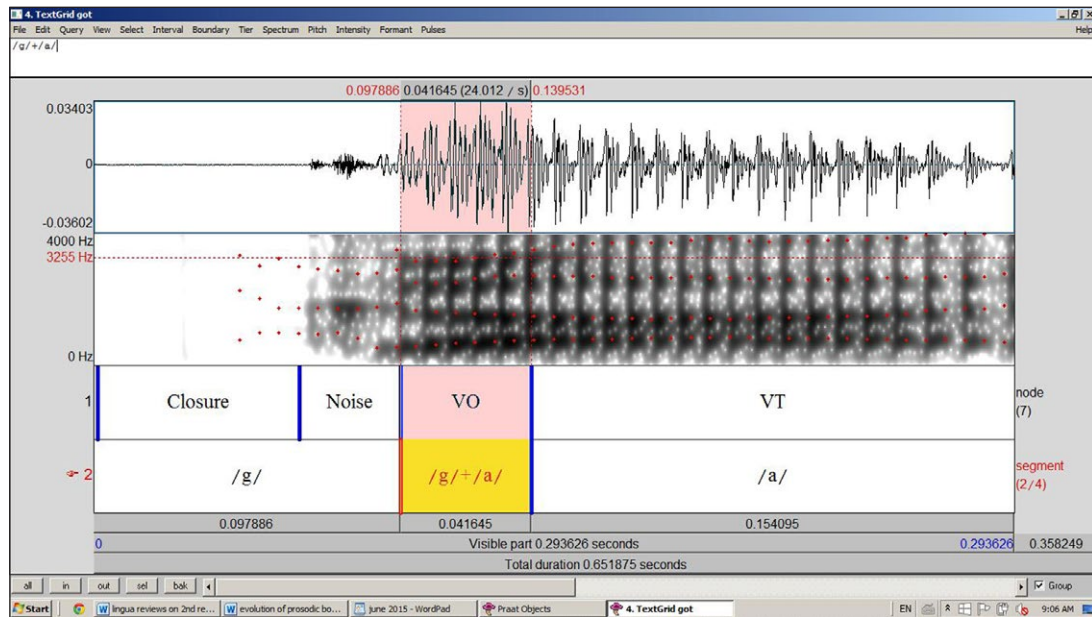
Figure 1: Sequence of OP structural nodes projected onto a CV sequence.

can only produce the sequence of acoustic events delineated in Figure 1. There is simply no other possibility.

This causality is reinforced by the fact that an individual property associated with Closure, place of articulation, directly affects the spectral properties of Noise and VO (and even VT). In OP, place of articulation is encoded as privative melodic specifications that attach to the nodes of the tree. In stops these are assigned at the Closure level, and occupy the lower-level Noise and VO nodes by means of a "trickling" mechanism (Schwartz 2016a: 45). The interaction between the OP hierarchy and the trickling mechanism formalizes the phonetic causality between consonant place of articulation and its acoustic consequences.[2] Positing a linear string would not capture these relationships. This is not to say that there is no linearity in phonology. Instead, the claim is that linearization of the universal OP hierarchy occurs on a language-specific basis, when the CV is unpacked into Cs and Vs. For further discussion of this issue, see Schwartz (2016a), in which it is shown that there are different ways of unpacking the CV, with far-reaching empirical consequences.

Figure 1 also allows for a visualization of carrier modulation, which is observable in two areas, the amplitude domain and the spectral domain. Assuming after MT that the carrier is a schwa-like vocoid, clearly stop closure, in which we may observe silence, represents the greatest acoustic departure from the carrier. Stevens (2002) has also recognized the importance of stop closures for listeners in his model of lexical access. Closures represent robust acoustic landmarks that allow listeners to parse continuous speech into smaller units. The noise burst produced by stop release also creates a salient modulation of the carrier, albeit not quite as salient as stop closure. The least salient modulations are those dealing with formant frequencies, since they do not alter the basic acoustic shape of the carrier in terms of periodicity and the presence of formant structure, only minimally affecting the amplitude envelope. Thus, in Figure 1, it is far less straightforward to identify formant modulations than amplitude modulations. At a point in the center of the CV transition that is highlighted in the spectrogram (aligning with the letters VO in

---

[2] Note that trickling moves progressively. Other mechanisms are required to show regressive effects of place of articulation. See Schwartz (2016a) for details.

the annotation), the formants resemble those of a schwa. Earlier in the vowel we see a lower F1 and higher F2, while later we see a lower F2 and a higher F1. Notice that unlike stop closures and noise, it is not entirely clear when looking at the vowel where formant modulations begin and end, leaving only the bare carrier.

Implicit in this discussion is a suggestion with regard to the fact that amplitude modulations are more easily identifiable in acoustic displays than spectral modulations. We claim that the visual robustness of manner-induced amplitude modulation in waveform displays is reflected in speech perception, such that manner contrasts are perceptually more salient, and more conducive to categorical perception than place contrasts (for more discussion, see Schwartz 2014). Consequently, it is amplitude modulations and manner of articulation that should comprise the building blocks of prosodic structure, with the most salient modulations (stop closure and aperiodic noise) occupying the highest positions in the OP representational hierarchy. In essence, the amplitude modulations associated with closure and noise represent quasi-discrete landmarks in the speech signal (cf. Stevens 2002).[3] Since phonology deals with discrete units, those phonetic events that are (more or less) discretely identifiable should be privileged with respect to phonological representation. In other words, since stop closures and aperiodic noise give structure to the acoustic signal, they are also the primary building blocks of prosodic structure. Meanwhile, spectral modulations are encoded as melodic specifications that attach to that structure. Thus, the dichotomy between amplitude and spectral modulation captures acoustic properties inherent to manner and place features, respectively.

The presence or absence of modulations in different types of CV sequences defines manner of articulation in segmental representations, given in (3), which are extracted from the OP hierarchy. Manner of articulation is encoded as the active (binary) nodes in a given structure, while the unary nodes act as placeholders that reflect missing phonetic events with respect to the entire CV hierarchy.[4] The segmental symbols are shorthand for place and laryngeal specifications to be developed in more detail shortly.

(3)   Individual segmental structures extracted from the OP hierarchy
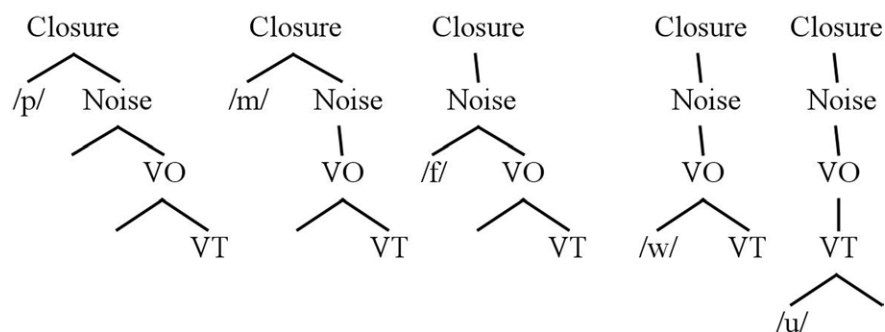
An important aspect of the trees in (3) is that segments do not link to "timing slots" that attach to prosodic structure. Rather, segments *are* prosodic structure, and different manners of articulation have different structural configurations. In many cases in which headedness has been evoked, the structural configurations of the OP environment,

---

[3] A reviewer raises the question of the relative discreteness of modulations associated with different types of consonant types. In particular, fortis or voiceless consonants are sometimes associated with more abrubt boundaries between acoustic events, particularly in the case of formant transitions (e.g. Stevens and Klatt 1974).

[4] A reviewer raises a question about the status of unary nodes, and whether they may be absent from the representations. The short answer to this question is yes. Empirical implications of the ambiguity between unary and absent nodes may be observed in areas such as sonority sequencing in phonotactics. For details, see Schwartz (2016).

which are derived directly from independently observable properties in the speech signal, allow us to encode relevant generalizations without requiring formal status for heads.
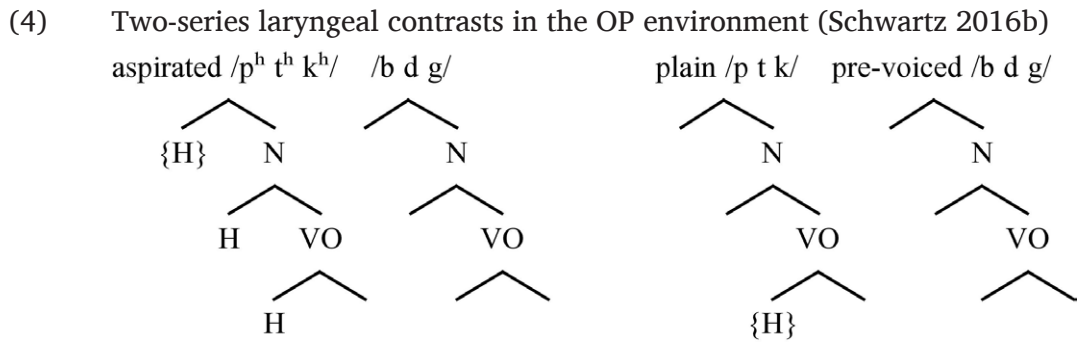
3 The voiced carrier and implications for laryngeal phonology

In Traunmüller's (1994) formulation of Modulation Theory, the carrier is assumed to be a periodic signal with evenly spaced formants that correspond to schwa. In principle, the carrier need not be voiced – aperiodic noise produced at the glottis acts as a carrier of whispered speech. However, such cases are clearly exceptional. A periodic signal makes a much better carrier given its robustness in the face of background noise that is present in most communicative situations (see e.g. Wright 2001).

These considerations have important implications for the phonology of laryngeal features. If the carrier signal, which is assumed to be lacking in phonological content, is voiced, it follows that voiceless obstruents represent more salient carrier modulations than voiced ones. As a result, the prediction of MT is that in languages with two series of obstruents, it is always the voiceless one that is phonologically specified, regardless of the realization of laryngeal contrasts in terms of pre-voicing or aspiration. This prediction is clearly at odds with earlier traditions, by which either [-voice] is the default value, or the choice of the specified variant is a function of the VOT contrast in the language (i.e. "Laryngeal Realism"; Honeybone 2005, Beckman et al. 2013). Of these traditions, the Laryngeal Realism approach has gained widespread following in recent years, since it directly encodes a measurable phonetic property, and appears to capture the generalization that plain voiceless unaspirated stops, which are claimed to be unspecified, appear to be typologically the most common type of obstruent (but see Vaux & Samuels 2005). In sum, there is a clear conflict between Modulation Theory and Laryngeal Realism as to the organization of voice contrasts. For Modulation Theory, voiceless should always be marked. For Laryngeal Realism, voiceless is only marked in aspiration languages.[5]

Onset Prominence representations allow us to reconcile this conflict such that voiceless is always phonologically specified as predicted by Modulation Theory, but the VOT typology underlying Laryngeal Realism may still be captured. The key aspect of OP representations for this purpose is that obstruents contain different levels of structure at which the single laryngeal specification (call it an element {H}) may be assigned. The level at which the laryngeal modulation appears determines the realization of the laryngeal contrast is terms of full voicing or aspiration. This is shown in (4). The two structures on the left obtain in aspiration languages, in which the {H} element is assigned at the Closure level and trickles down to occupy Noise and VO. The pair of structures on the right represent laryngeal contrasts in languages with fully voiced obstruents, in which the {H} specification is assigned at the VO level, leaving Closure and Noise unspecified.

---

[5] In this paper, the term "marked" is used interchangeably with "specified", and "unmarked" with "unspecified". No claims are made with respect to "markedness" as an indicator of typological rarity or difficulty in acquisition. The traditional assumption is that "marked" features are acquired later, and are less common. However, this view of markedness is not universally accepted (Blevins 2004; Haspelmath 2006), and is full of contradictions. For example, Hume (2011) notes that both "perceptually weak" and "perceptually strong" phonetic properties have been claimed by various authors to be unmarked, and marked. In this connection, my claim that voiced is unspecified should not be interpreted in terms of these traditional views on markedness.

(4)   Two-series laryngeal contrasts in the OP environment (Schwartz 2016b)

aspirated /pʰ tʰ kʰ/   /b d g/   plain /p t k/   pre-voiced /b d g/

{H}   N   N   N   N

H   VO   VO   VO   VO

H   {H}

Two aspects of the representations in (4) require further comment at this time. First and foremost, the possibilities for {H}-assignment provide a phonetically faithful perspective of the difference between voiceless aspirated and plain voiceless stops.[6] The lack of aspiration in the plain stops is reflected by the fact that the Noise node contains no laryngeal specification, while aspiration is indicated by H on Noise. The other thing to notice is that pre-voiced and unvoiced lenis stops have the same representation regardless of laryngeal type. Thus, pre-voicing, when it appears, does not reflect phonological specification. Rather it may be thought of as part of the carrier, with the effect that Closure and Noise modulations are weakened.

Both phonological arguments and phonetic facts may be marshalled to support this outlook on pre-voicing in voicing languages. On the phonological side, Cyran (2013) and van der Hulst (2015) have shown that privative systems explain voicing assimilation facts from dialectal Polish and Dutch, respectively, only under the assumption that voiced is unspecified. Other authors have observed the phonological activity of voicelessness in "voice" languages to argue for binary specification of the feature [voice]. For example, Rubach (1996) cites progressive devoicing processes in Polish (e.g. of <rz> in words such as *przy* [pʂɨ] 'by') as a case in which voicelessness in a true-voice language may spread, and claims that [-voice] must be phonologically active. Additionally, Wetzels & Mascaró (2001) show that voicelessness may spread in many other true-voice languages, including Romanian, French, and Bosnian/Croatian/Serbian, and argue for a binary approach to laryngeal specifications on phonological grounds. These apparent binary effects are captured in the representations in (4), despite the fact that all the specifications are monovalent.

On the phonetic side, it has been shown that laryngeal contrasts are robust in voicing languages even in the absence of closure voicing. For example, in Dutch (van Alphen & Smits 2004) and Afrikaans (Coetzee et al. 2014), there is a great deal of variability with regard to the appearance of pre-voicing in the "voiced" set of obstruents. However, the laryngeal contrasts in these languages are maintained on the basis of other cues. Indeed, the Afrikaans case has been described as in instance of emergent tonogenesis, by which f0 appears to be taking over from VOT as the primary cue to the laryngeal contrast. Polish also shows variability in pre-voicing, yet the laryngeal contrast is robustly perceived even in its absence (Schwartz et al. 2017). Additional evidence comes from the speech Polish learners of English, who are typically more successful in acquiring English-style aspiration (Zając 2015; Schwartz et al. 2017) than they are in producing native-like lenis consonants

---

[6] In fact, there are three different possibilities for {H} assignment: Closure, Noise, and VO. Closure-level *vs.* Noise-level VO assignment could be used to encode the difference between aspiration languages such as Icelandic with pre-aspiration and those without. Developing this analysis is a task for future work.

without voicing. These results suggest that the aspirated stops are "new" and more easily acquired in accordance with Flege's (1995) Speech Learning Model, while the L2 lenis stops are subject to "equivalence classification" (Flege 1987) and confused with L1 /bdg/. This interpretation of the L2 facts is compatible with the representations in (4), in which /bdg/ are representationally identical in voice and aspiration languages. However, LR incorrectly predicts that both lenis *and* fortis English stops should be new to speakers of voicing languages, since both have different representations in L1 and L2.[7]

So far we have seen how the predictions of MT with regard to manner and laryngeal specifications may be encoded with Onset Prominence representations. At this time, we turn our attention to a discussion of the relationship between nasality and voicing, which many authors have attributed to headedness. It will be shown that no headedness device is necessary to capture this relationship.

3.1 Nasality and voicing

Linguists have long noted that there is a connection between voicing and nasality. This link is a clear instantiation of phonetic effects that may be observed in gradient form or generalized to attain the status of phonological generality in many languages (see Hayes 1999). The phonetic side of the equation rests on two facts. First, since phonation requires continuous airflow through the glottis, which can be hindered by a supra-laryngeal constriction, lowering the velum provides a passage for continued airflow and thus facilitates voicing. Second, producing a sequence of nasal + voiceless stop (NT) is relatively difficult as it requires the coordination of velum opening with the release of the oral closure.

Beyond these phonetic considerations, a number of phonological generalizations note the connection between nasality and voicing. First and foremost, nasal consonants are almost always voiced. Only a small percentage of the languages of the world feature nasals that are unvoiced (or produced with non-modal phonation), and these languages always have modally voiced nasals as well. For example, in the UPSID database (Maddieson 1984), /m/ occurs in approximately 95% of the languages included, while the voiceless bilabial nasal appears in under 4% of the languages. In addition, Ladefoged & Maddieson (1996) note that voiceless nasals are typically partially voiced near the release of the oral constriction, often giving the auditory impression of /hn/ sequences. Indeed, in some cases the presence of voiceless nasals in a language has been shown to derive from consonant sequences (see Botma 2004). Another generalization concerns (NT) clusters, which are prohibited in many languages. In morphological environments where NT clusters might form, the stop is typically voiced.

In Element Theory, the nasal-voicing connection has been expressed as a claim that nasality and voicing in obstruents reflect the presence of a single monovalent primitive (Nasukawa 1998, 2005; Botma 2004; Breit 2013), usually the element {L} (Nasukawa uses the symbol {N}), reflecting the low frequency periodicity of the voice bar and nasal resonance. Under this view, the representational distinction between voicing in obstruents and nasals is due to headedness. Interestingly, Element Theorists do not agree as to whether nasals or voiced obstruents contain the headed version of the L element.

---

[7] Processes of intervocalic voicing constitute another problem for Laryngeal Realism. According to Beckman et al. (2013), LR predicts that intervocalic voicing should only be attested in aspiration languages, and absent from true-voice languages. In aspiration languages, the process is straightforward for LR – the [SG] specification is lost in a weak position. However, for a LR description of intervocalic voicing in true-voice languages, it would be necessary to insert a [voice] specification on a consonant that lacks one. It is therefore predicted to be impossible. As it turns out, however, intervocalic voicing is quite common in true-voice languages (see e.g. Hualde & Nadeu 2011 for Rome Italian; Hualde et al. 2011 for Spanish; Keating 1980 for Polish), yet such intervocalic voiced tokens tend to be perceived as voiceless.

Nasukawa argues that voiced obstruents contain headed {L}, while Breit (2013) argues that headed {L} is found in nasals.

OP representations capture the voicing-nasality relationship without the need to appeal to headedness. Consequently, any disagreement as to the type of consonant associated with headed {L} is rendered moot. Stops and nasals are structurally distinct, and the presence of voicing in nasals is simply the periodicity inherent in the carrier. In the OP environment, the fundamental difference between a stop and a nasal lies in the status of the Noise node, which is active only in the former, encoding aperiodic release bursts that are absent in nasals. This is illustrated in the structures in (5), with the stop on the left and the nasal on the right.

(5)   Stops vs Nasals in the OP environment
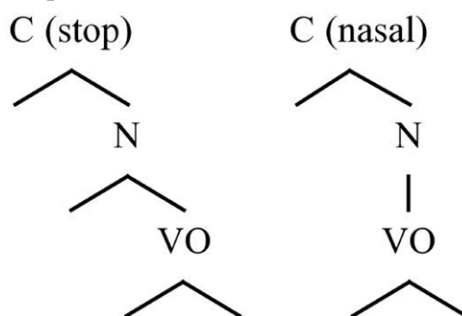
When the laryngeal contrasts outlined in (4) are projected on these structures, predictions are made with regard to the interaction between nasality and voicing, in particular the possibility of laryngeal contrast in nasals.

Consider the structures in (6), in which we see labial stops shown in the two types of language. In aspiration languages shown on the left, both place and laryngeal specifications are assigned at the Closure level and "trickle" down the consonantal structure to occupy the Noise and VO nodes. In voicing languages, place is assigned at Closure and the laryngeal specification is assigned at VO.

(6)   Place and laryngeal specifications in aspiration vs. voicing languages
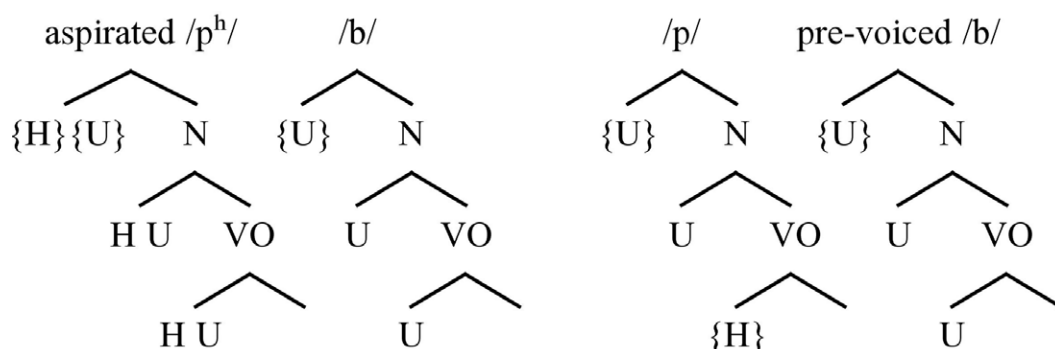

The trickling mechanism serves to provide a given specification with more structural housing for phonetic realization, and thus greater perceptibility. In voicing languages, in which place and laryngeal specifications are assigned at different levels (Closure for place, VO for laryngeal), the configuration suggests a restriction against trickled and assigned specification on the same node. That is, the assignment of {H} to VO in unaspirated /p/ (3rd tree from the left) blocks the trickling of the place specification. This restriction is formalized as a constraint, BLOCKTRICKLING, given in (7).

(7)   BLOCKTRICKLING – Trickling of a melodic specification is blocked by the
        assignment of an additional specification at a lower level.

Note that this constraint only affects place trickling in voicing languages. In aspiration languages, both place and laryngeal specifications are assigned at the same level, so both are free to trickle until they reach melody associated with the following vowel. Some phonetic implications of this proposal, particularly with regard to the relative weight of VO-level CV transitions for place perception, will be discussed momentarily. At the moment however, we must turn our attention to nasals.

Nasal representations are given in (8). The structure on the left is a regular voiced nasal in both voicing and aspiration languages. The middle tree is a voiceless nasal in an aspiration language. The rightmost tree shows what a voiceless nasal would look like in a voicing language.

(8)   Voiced nasals, voiceless nasals, and a hypothetical voiceless nasal in a voicing
        language
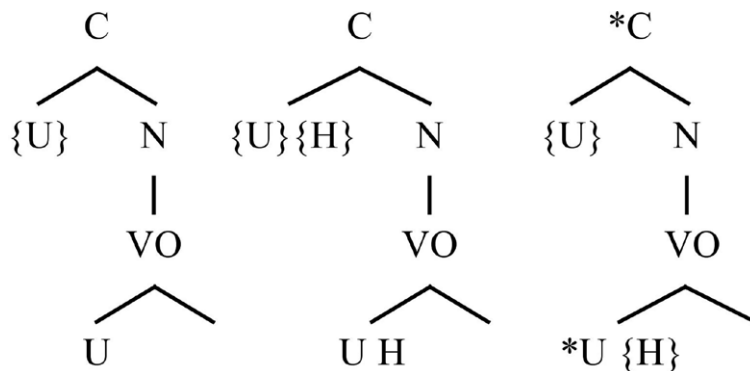
Since nasals lack the Noise node, we should expect a prohibition on voiceless nasals in voicing languages. If, as shown in the structure on the right, the {H} specification blocks the trickling of place onto the VO node, only the Closure level is available for the phonetic expression of the place feature, rendering that feature essentially inaudible. Thus, voiceless nasals should only be allowed in languages that allow VO to be occupied by both trickled place and laryngeal specifications. Even in those cases, however, we should expect voiceless nasals to be rare, since they typically require a period of voicing just before the release of the oral constriction (Ladefoged & Maddieson 1996), yielding the auditory impression of a two-segment sequence /hn/.

An additional prediction that falls out from this discussion is that categorical (as opposed to gradient, cf. Hayes 1999) post-nasal voicing should be restricted to voicing languages. Unaspirated stops have a Closure node that is unspecified for laryngeal features, so these stops may undergo full voicing without the loss of a melodic specification. By contrast, when the Closure node is specified with {H}, some degree of voicing may occur, but it should be gradient in nature, and not sufficient to neutralize the laryngeal contrast.[8] Whether these predictions are borne out is an empirical question that I plan to take up in the near future. However, it is not unreasonable to expect that the phonological status of a

---

[8] A reviewer asks about non-foot-initial fortis stops in English, suggesting that since they are often described as unaspirated, they should lack Closure-level {H} specification. This is not my prediction at all. Rather, these are fortis stops that are still clearly distinct from lenis stops, but with weaker aspiration. Because they are in a weaker position, they are subject to gradient post-nasal voicing as described by Hayes (1999).

post-nasal voicing process should be dependent more on the nature of the voicing contrast than on the nasal.

We have seen that the OP perspective on the relationship between voicing and nasality makes insightful predictions with regard to the appearance of laryngeal contrasts in nasals, and the question of whether post-nasal voicing is gradient or categorical. These predictions are also expressible in Element Theory approaches to these phenomena, so I must say a few words about why the OP approach is preferable. Stated briefly, the configurations outlined above derived from independently motivated phonetic facts, while finding independent motivation for headedness is a much more difficult endeavor. In what follows, we shall briefly review this claim.

Crucial for the OP account of the nasal-voicing connection is the claim that place specifications are obligatorily assigned at the Closure level, but laryngeal specifications may be assigned at the Closure or VO level, depending on the language. The phonetic connection between place and Closure is obvious – the location of a constriction is the defining property of a place specification. Thus, the correspondence between place assignment and Closure is to be expected. By contrast, languages show a great deal more variety in the timing of laryngeal events with regard to consonant constrictions (e.g. Ladefoged & Maddieson 1996). However, OP allows us to restrict this variability and make it phonologically manageable. The assignment of {H} to VO in voicing languages is a natural expression of voiceless unaspirated stops, since the higher-level Noise node is unaffected by the laryngeal feature. When {H} is assigned to Closure, it trickles down to occupy the Noise node, which provides a prosodic docking point realized as /h/-like noise.

The opposition between languages with VO-level laryngeal assignment and those with Closure-level laryngeal assignment leads to important predictions for cross-language phonetic study. In voicing languages, laryngeal assignment blocks the trickling of the place specification from the VO node associated with the CV transition. Therefore, the prediction is that since Noise is the only node available for place realization, stop release bursts should be a more robust place cue in voicing languages than in aspiration languages, in which formant transitions should carry greater perceptual weight. Evidence for this prediction may be found in an experimental study carried out by Schwartz & Aperliński (2014), who used cross-spliced stimuli to test the relative weight of noise bursts vs. formant transitions for stop place perception in Polish and English. In Polish, a voicing language, Noise bursts were weighted more heavily than in English, in which formant transitions were dominant (cf. Walley & Carrell 1983). These predictions follow naturally from the OP representation of voicing, but are not expressible in standard element theory.

To conclude this section, we have seen that the relationship between voicing and nasality, traditionally described in Element Theory in terms of headedness, falls out directly from independently motivated aspects of OP representations. There is no need for headedness as a formal device. At this point we turn our attention to vowel quality, in which harmony patterns in some languages exhibit asymmetries that are suggestive of headedness.

4 Spectral modulation and the auditory anatomy of place elements

In the trees in (3), the segmental symbols may be thought of as shorthand for a single labial specification, which we refer to after element theory (Harris & Lindsey 1995) as {U}. Perceptual cues associated with this {U} specification vary as a function of the level of the OP hierarchy. At the Closure level, place cues are largely inaudible unless the Closure is preceded by a vowel and VC transition, a falling F2 in the case of labials. Labial noise is low in amplitude and characterized by a relatively flat spectrum. On VO, labials are associated with a rising F2 of the CV transition. Finally, under the VT level, /u/-like vowel quality is known to be characterized by a low second formant and a low first formant.

What follows from these facts is that the acoustic signature of {U} is dependent on the level of the OP hierarchy at which the element appears. Since Element Theory claims that there is direct mapping between phonological representation and speech without a level of "categorical" phonetics (cf. Harris 2004), and that elements are defined in terms of their acoustic properties, any element-based implementation of the phonetics-phonology interface must allow for the possibility that elements may be broken down into smaller parts. That is, F2 rises or falls, flat noise spectra, and low formant targets may all constitute building blocks in the auditory structure of {U}. In what follows, we will develop this idea further with regard to vowel quality, exploiting a proposal for the internal anatomy of elements outlined in Schwartz (2009).

In Harris & Lindsey (1995), elements are presented as independently interpretable primes, that if translated into traditional feature theory would constitute amalgams of several features specified with binary values.[9] Thus, the ET approach combines what feature theory would claim are three specifications, [+high] [+back], and [+round], into a single primitive {U} that is interpretable as the vowel /u/. What earlier presentations of Element Theory do not consider is the possibility that the phonetic building blocks of privative elements are privative themselves. This claim is more or less explicit in Modulation Theory – a modulation is present if it is salient enough to be perceptible, otherwise it is absent. In this connection, speech perception research has shown that acoustic cues to vowel features should be seen in privative terms. While phonetics textbooks note the correlations between F1, F2 and traditional vowel charts, a number of studies have shown that vocalic feature categories are also perceived in terms of spectral convergences. For example, it is not only a high F2 that cues front vowels, but also the perceptual convergence of F2 with F3 (Syrdal & Gopal 1986). Likewise, the phonological feature [+high] is perceived when F1 is converged with f0 (Hoemeke & Diehl 1994). Auditory experiments have established that convergence of two spectral prominences takes place when two formants are within 3 Bark of each other (Chistovich et al. 1979).[10]

So far, we have discussed spectral convergences as cues to front vowels (F3–F2) and high vowels (F1–f0). However, representing vowels solely in terms of convergences says nothing about the actual formant frequencies of vowels with respect to the schwa-like carrier. Due to the non-linear relationship between the psychoacoustic Bark scale and the purely acoustic Hertz scale, we can envision a scenario in which a formant may be raised or lowered to converge with another spectral prominence, yet still be indistinguishable from the baseline value for schwa. Alternatively, a formant may be clearly distinct from the baseline, yet not be converged with another spectral prominence. It follows therefore, that any attempt to incorporate spectral modulations into a model of phonological primitives must consider two types of modulation: convergence cues *and* cues associated with single formants. In this connection, we should expect the auditory thresholds relevant for the percept of spectral convergences to differ from those of a single formant relative to schwa. That is, a single formant modulation is a different perceptual creature from a modulation that combines two formants. The Bark unit is an auditory critical band, so any single formant within 1 Bark of the schwa baseline should be perceived as schwa-like. At

---

[9] Harris & Lindsey (1995) explicitly state that such "translation" does not occur: specifications like [+high], [–back] do not exist in ET

[10] Strictly speaking, f0 is not a "formant" since it is not directly related to vocal tract resonance properties. However, considering the typical spectral characteristics of phonation in which the fundamental is the strongest harmonic and the amplitude of the subsequent harmonics decreases (except perhaps in the case of stiff or creaky voice qualities; Gordon & Ladefoged 2001), it is reasonable to label f0 as a spectral "prominence".

the same time, as mentioned earlier, two formants need only be less than 3 Bark apart to be perceived as converged (Chistovich et al. 1979).

Let us consider some examples in which converting acoustic measurements to the Bark scale results in mismatches between formant convergences and single formants. Table 1 includes mean F1 and f0 values of American English tense and lax high vowels (/i ɪ u ʊ/) for male speakers from the study by Hillenbrand et al. (1999). The original Hertz values are converted into the Bark scale according to Traunmüller's transformation (Traunmüller 1990).[11]

The measurements in the table show that the lax vowels are characterized by the F1f0 convergence associated with high vowels, but they lack an F1 that is more than 1 Bark away from the schwa baseline. Thus, the "headed" nature of the {I} element in the tense vowel is expressible simply as the presence or absence of a privative LowF1 cue. For an illustration of the anatomy of {I} and {U} in more detail, consider Tables 2 and 3. The most basic component of {I}, appearing in all front vowels, is the F3F2 convergence. Meanwhile the most basic component of {U} is the LowF2 cue. The fact that for {I} the fundamental building block is a formant convergence, while for {U} it is a single formant cue, will become relevant in our discussion of vowel harmony

To conclude this section, it may be noted that headedness effects in vowel quality are derivable from measurable auditory properties. To visualize this it is necessary to adopt a perspective, suggested by Modulation Theory, from which perceptual cues to phonological primes are privative in nature. With this strategy, we can express the basic assumption

Table 1: Bark values for American English high vowels (transformed from Hillenbrand et al. 1999).

Vowel	F1 schwa	f0	F1	F1 more than 1 Bark from schwa?	Convergence w f0?
/i/	4.9	1.4	3.5	Yes	Yes
/ɪ/	4.9	1.3	4.2	No	Yes
/u/	4.9	1.4	3.8	Yes	Yes
/ʊ/	4.9	1.4	4.4	No	Yes

Table 2: Internal structure of {I} element for front vowels.

Vowel	{I} cues
i	F3F2, F1f0, HighF2, LowF1
ɪ	F3F2, F1f0, HighF2
e	F3F2, HighF2
æ	F3F2, HighF2, (HighF1)

Table 3: Internal structure of {U} element for back vowels.

Vowel	{U} cues
u	LowF2, F1f0, LowF1
ʊ	LowF2, F1f0
o	LowF2

---

[11] The Bark scale is an auditory transform that translates the human hearing range (up to 20 kHz) into 24 critical bands. Traunmüller's equation is as follows: Bark = ((26.81 * Hz)/(1960 + Hz))–0.53; if Bark < 2, add 0.15 * (2-result)

of Element Theory that phonological representations map directly to the speech signal. At the same time, it is important to note that this approach to the phonology of vowels is predictive of the variability that is the focus of many phonetic and sociolinguistic studies. The privative formant cues proposed here are all relative in nature; formants are expected to fit into certain acoustic windows, but no claims are made about the exact location of those windows on the frequency scale – actual formant frequencies are of course subject to a great deal of variability.

4.1 Asymmetries in vowel harmony

So far, we have seen how two types of spectral modulations may constitute building blocks of place elements in the representation of vowel quality. Formant convergences are posited when two spectral prominences come within 3 Bark of one another, while single formants may constitute salient modulations when they stray more than 1 Bark from a baseline value associated with schwa. In a sketch of the internal structure of elements used to represent vowel quality, it was proposed that the relative salience of a single element in a given vowel is derived not from headedness, but from the number of privative modulations found in the elemental structure.

One asymmetry between the elements {I} and {U} was noted in this discussion. The most basic modulation associated with the element {I}, present in all front vowels, is the convergence of the second and third formants (F3F2). Meanwhile, the only modulation present in all instances of the element {U} is a low second formant (LowF2). Thus, there is a fundamental difference in the auditory structure of the two elements in that {I} is built on a formant convergence and {U} is built on a single formant. Consequently, we might expect asymmetries between the behavior of {I} and {U}. Formant convergence modulations, since they occupy a larger part of the spectrum, should be more perceptually robust than single formants. Thus, we should expect {I} to be a better candidate for harmony than {U}.

Just such an asymmetry has been observed in vowel harmony patterns. Charette & Göksel (1996) describe vowel harmony in several Turkic languages, including Turkish, Yakut, Kazakh and Kyrgyz. In each of these languages, there are cases when the element {I} is allowed to spread but the element {U} is not. An example from Turkish is given in (9).

(9)   *Turkish* (Charette & Göksel 1996: 13)
        Stem   Dative
        kuʃ   kuʃ-ta   *kuʃ-to   'bird'
        jyk   jyk-te   'load'

In the first case, the vowel in the dative suffix –*ta* is unaltered, the {U} that is present in the stem vowel does not spread. In the second example, conversely, the {I} from the stem is spread but the {U} is not, yielding a front but unrounded vowel /e/ in the dative suffix. Charette & Göksel (1996) invoke headedness in their explanation of this asymmetry, proposing a constraint that in complex expressions, {U} must be the head, and is therefore blocked from spreading into suffixes containing headed {A̲}. The element {I} is not subject to this restriction and is free to spread. With the representations here, there is no need to invoke headedness – the internal structure of {I} and {U} – one based on a formant convergence the other on a single formant – explain why the two elements behave differently.

The basic picture that emerges from the Turkish examples, as well as other cases from Turkic languages discussed by Charette and Göksel, including Yakut, Kazakh, and Kirghyz, is that rounding harmony is disadvantaged with respect to palatal harmony. Rhodes (2010) notes that rounding harmony is found in far fewer languages than palatal

harmony. In addition, when rounding harmony is attested, it is always accompanied by another kind of harmony, which is typically more widespread in the language.

One non-Turkic example of this type is Akan (O'Keefe 2003), in which there is widespread ATR harmony, but rounding harmony is limited to certain affixes on a dialectal basis. From the point of view of Modulation Theory, tongue root advancement, like palatality, is strongly associated with a formant convergence, in this case F1f0.[12] As such, it may be assumed to be show similar harmony asymmetries with respect to rounding, which is built on a single formant cue (Low F2). The data from Akan show that this is the case. In (10) we see examples of stems with two or more syllables. They all agree in ATR, but not in rounding.[13]

(10)   *Akan* ATR agreement within stems (O'Keefe 2003)
    a.   With rounding agreement
        efie   'home' ɛfɪɛ   'vomit'
        owu   'death' ɔwʊ   'snake'
        koko 'chest' kɔkɔ 'crab'

    b.   Without rounding agreement
        bosome 'month/moon' ɛtʊrɔ 'lie'
        bɔnɪ   'evil'   kube 'coconut'

In affixed forms, Akan rounding harmony is found with certain affixes in given dialects of the language. Examples in (11) include the progressive prefix /rV[+high]/ in (10a), and the ingressive prefix (be/bɛ/bo/bɔ), as in (11b) and (11c).

(11)   Dialect-specific rounding harmony in *Akan* (O'Keefe 2003)
    a.   ɔ.rɪ.kɔ   ɔ.rʊ.kɔ (*Fante* dialect) 'he is going'
    b.   o.be.tu   o.bo.tu (*Fante* dialect) 'he comes and digs up'
    c.   ɔ.b.ɛ.kʊ   ɔ.bɔ.kʊ (*Fante* dialect) 'he comes and fights'

Thus, in Akan, as in Turkic, it appears that rounding harmony is something of a marginal phenomenon relative to the other type of harmony appearing in the language. Indeed, O'Keefe notes that the earliest descriptions of the language make no reference to rounding harmony at all.

The disadvantaged status of rounding harmony from the point of view of Modulation Theory may also be reflected in height restrictions found in rounding harmony languages, which may serve to compensate for the acoustic handicap of rounding harmony. Examples are discussed by Kaun (1995), who notes that rounding is more likely to spread from non-high vowels than from high vowels. That is, /o/ is more like to trigger the {U} harmony than /u/. Some examples from Yakut (Turkic) and Burjat (Eastern Mongolic) are given in (12).

(12)   Rounding harmony asymmetries in Yakut and Burjat (Kaun 1995)
    a.   *Yakut* (Kaun 1995: 25–26)
        torbos-tor   'heifer' (PL)
        tunnuk-ter 'window' (PL) *tunnuk-tor

---

[12] ATR may ensure an F1f0 convergence in two ways. First of all, by lowering F1, it reduces the acoustic distance between F1 and f0 frequencies. In addition, it has been shown that ATR increases the bandwidth of F1 (Hess 1992), essentially widening the spectral range of the formant to absorb f0.

[13] Tungusic languages have been claimed to exhibit retracted tongue root (RTR) harmony (Ko 2012). From the point of view of MT, RTR might be encoded as a convergence between F1 (which is raised) and F2, or as phonetic prominence associated with higher spectral tilt.

   b.   *Buriat* (Kaun 1995: 47)
        to:n-do: 'white spot' (DAT)
        xul-de: 'foot' (DAT) *xul-do:

In these examples, the suffixes (*-ter* for Yakut, *-de:* for Buriat) contain a mid vowel that is unchanged after /u/ in the stem. When the stem vowel is mid, however, it triggers {U} harmony, and the vowel surfaces as /o/ (or /o:/).

Kaun (2004) attributes such patterns to perceptual factors, suggesting that the function of rounding harmony is to render the rounded vowel quality more salient to listeners. The degree of rounding and its acoustic effects are less robust on mid vowels than on high vowels (cf. Donegan 1985), so spreading increases the chances that the rounding will be perceived. In other words, the {U} element on a mid vowel needs "help" in order to be heard, and thus spreads. From the point of view of Modulation theory, high vowels show more robust realizations of the {U} element than mid-vowels. While both contain a Low F2 modulation (a single formant cue), the mid vowel lacks the F1f0 convergence that contributes to the realization of {U} in high vowels. The lack of the convergence constitutes the motivation for spreading, which increases the chance that the {U} will be perceived by the listener.

To summarize the harmony discussion, we have seen how the Modulation-based view of the anatomy of elements may explain how (1) rounding is disadvantaged with respect to other harmonic features such as palatality and tongue root advancement, and (2) how restrictions on rounding harmony fall out from the inherent modulatory properties of rounding. In accounting for the Turkic data exemplifying the first of these patterns, Charette and Göksel (1996) propose that {U} is subject to a licensing constraint but {I} is not. While their explanation encodes the harmony asymmetry between the two elements, it is essentially arbitrary. There is nothing to prevent another language from imposing a constraint on {I}, while allowing unrestricted spreading of {U}. To my knowledge, no such language exists. By contrast, the Modulation account shows that these asymmetries have independent phonetic motivation.

To conclude the discussion on the internal anatomy of the elements, it is necessary to discuss the implications of this perspective for the compatibility or lack thereof between the Modulation approach and traditional ET. At first glance, it appears as if we are breaking with Element Theory in suggesting that phonological patterns may be explained in terms of primitives that are smaller than elements. This would indeed be the case if we were to assume that elements are universal primitives of phonological representation. The Onset Prominence theory, however, does not make this assumption. Melodic features in OP are language-specific emergent primes which often behave differently in different languages. The Modulation perspective aims to offer a story about how such differences may be explained from an evolutionary perspective. What is in line with the fundamental epistemology of ET is our assumption that these cues are privative. Monovalent primitives are inherently categorical, and thus phonological, in nature. In this way, our approach contrasts with other "phonetically-based" approaches that focus their attention on the gradient properties of speech.

5 Conclusion

In this paper I have argued that in the domain of segmental representation, effects that in Element Theory have been attributed to headedness are in fact derivable from more general phonetic considerations. The representations of the Onset Prominence framework provide a perspective from which the origins of these effects may be explained. OP representations encode the assumptions of Modulation Theory, according to which phonological specifications are built from salient modulations to a carrier signal. These

modulations are privative and thus lend themselves to phonological interpretation. That is, they offer primitive building blocks from which we can describe a direct mapping between phonological representation and the speech signal (cf. Harris 2004).

In the case of manner of articulation, modulations affect the amplitude envelope, which provides the building blocks of prosodic structure in the OP environment. With regard to voicing, the periodic nature of carrier requires that voiced is the unmarked laryngeal specification in obstruents, and the well-known VOT typology (voicing vs. aspiration languages) is derived from the level of the OP hierarchy at which the element {H} is assigned. The interaction of manner and laryngeal phonology in the OP environment is capable of capturing the oft-described relationship between nasality and voicing without recourse to headedness. Finally, with regard to place elements in vowel quality, modulations based on both formant convergences and single formant frequencies provide building blocks of the elements {I} and {U}. The internal structure of these elements allow for an explanation of asymmetries in vowel harmony without the need for headedness as a formal device.

Abbreviations

DAT = dative, ET = Element Theory, MT = Modulation Theory, OP = Onset Prominence, PL = plural, VO = Vocalic Onset, VT = Vocalic Target

Acknowledgements

Competing Interests

The author has no competing interests to declare.

References

Anderson, John & Colin Ewen. 1987. *Principles of dependency phonology*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511753442

Beckman, Jill, Michael Essen & Catherine Ringen. 2013. Evidence for laryngeal features: Aspirating vs. true-voice languages. *Journal of Linguistics* 49(2). 259–284. DOI: https://doi.org/10.1017/S0022226712000424

Blevins, Juliette. 2004. *Evolutionary Phonology – the emergence of sound patterns*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511486357

Botma, Bert. 2004. *Phonological aspects of nasality – an element-based dependency approach*. Amsterdam: University of Amsterdam dissertation.

Breit, Florian. 2013. Voice-nasal interaction in headedness in voiceless nasals. *UCL Working papers in linguistics* 25. 1–21.

Charette, Monik & Asli Göksel. 1996. Licensing constraints and vowel harmony in Turkic languages. *SOAS Working Papers in Linguistics* 6. 1–25.

Chistovich, L. A., R. I. Sheikin & V. V. Lublinskaya. 1979. Centres of gravity and spectral peaks as the determinants of vowel quality. In Bjorn Lindblom & S. Ohman (eds.), *Frontiers of speech communication research*, 143–157. London: Academic Press.

Coetzee, Andreas, Patrice Beddor & Daan Wissing 2014. Emergent tonogenesis in Afrikaans. *Journal of the Acoustical Society of America* 135. 2421. https://doi.org/10.1121/1.4878043

Cyran, Eugeniusz. 2013. *Polish voicing – between phonology and phonetics*. Lublin: Wydawnictwo KUL.

Donegan, Patricia. 1985. *On the natural phonology of vowels*. New York: Garland.

Flege, James. 1987. The production of 'new' and 'similar' phones in a foreign language – evidence for the effect of equivalence classification. *Journal of Phonetics* 15. 47–65.

Flege, James. 1995. Second language speech learning – theory, findings, and problems. In Winnifred Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, 233–277. Timonium, MD: York Press.

Golston, Chris & Harry van der Hulst. 1999. Stricture is structure. In Ben Hermans & Marc van Oostendorp (eds.), *The derivational residue of Optimality Theory*, 153–173. Amsterdam: John Benjamins.

Gordon, Matthew & Peter Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29. 383–406. DOI: https://doi.org/10.1006/jpho.2001.0147

Harris, John. 2004. Release the captive coda – the foot as domain of phonological interpretation. In John Local, Richard Ogden & Rosalind Temple (eds.), *Phonetic interpretation – Papers in laboratory phonology VI*, 103–129. Cambridge: Cambridge University Press.

Harris, John & Geoff Lindsey. 1995. The elements of phonological representation. In Jacques Durand & Francis Katamba (eds.), *Frontiers of phonology – atoms, structures, derivations*, 34–79. London/New York: Longman.

Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 46. 25–70. DOI: https://doi.org/10.1017/S0022226705003683

Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and case studies*. Chicago, IL: University of Chicago Press.

Hayes, Bruce. 1999. Phonetics in phonology – the role of Optimality Theory and inductive grounding. In Michael Darnell, Edith Moravcsik, Michael Noonan, Frederick Newmeyer & Kathleen Wheatley (eds.), *Functionalism and formalism in linguistics* 1. 243–285. Amsterdam: John Benjamins.

Hess, Susan. 1992. Assimilatory effects in a vowel harmony system: An acoustic analysis of advanced tongue root in Akan. *Journal of Phonetics* 20. 475–492.

Hillenbrand, James, Laura Getty, Michael Clark & Kimberlee Wheeler. 1999. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97. 3099–3111. DOI: https://doi.org/10.1121/1.411872

Hoemeke, Kathryn & Randy Diehl. 1994. Perception of vowel height – the role of F1–f0 distance. *Journal of the Acoustical Society of America* 96. 661–674. DOI: https://doi.org/10.1121/1.410305

Honeybone, Patrick. 2005. Diachronic evidence in segmental phonology: The case of obstruent laryngeal specifications. In Marc van Oostendorp & Jeroen van de Weijer (eds.). *The internal organization of phonological segments*, 319–354. Berlin: Mouton de Gruyter.

Hualde, Jose & Marianna Nadeu. 2011. Lenition and phonemic overlap in Rome Italian. *Phonetica* 68. 215–242. DOI: https://doi.org/10.1159/000334303

Hualde, Jose, Miquel Simonet & Marianna Nadeu. 2011. Consonant lenition and phonological recategorization. *Laboratory Phonology* 2. 301–329. DOI: https://doi.org/10.1515/labphon.2011.011

Hume, Elizabeth. 2011. Markedness. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume & Keren Rice (eds.), *Companion to phonology*, 79–106. Oxford/Boston: Blackwell-Wiley.

Kaun, Abigail. 1995. *The typology of rounding harmony – an Optimality Theoretic approach*. Los Angeles, CA: UCLA dissertation.

Kaun, Abigail. 2004. The typology of rounding harmony. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically based phonology*, 87–116. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511486401.004

Ko, Seongyeon. 2012. *Tongue root harmony and vowel contrast in Northeast Asian languages*. Ithica, NY: Cornell University dissertation.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Cambridge: Blackwell.

Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511753459

Nasukawa, Kuniya. 1998. An integrated approach to nasality and voicing. In Eugeniusz Cyran (ed.), *Structure and interpretation — studies in phonology*, 205–225. Lublin: Folium.

Nasukawa, Kuniya. 2005. *A unified approach to nasality and voicing*. Berlin: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110910490

O'Keefe, Michael. 2003. Akan vowel harmony. Swarthmore, PA: Swarthmore College BA thesis.

Pöchtrager, Markus. 2006. *The structure of length*. Vienna: University of Vienna dissertation.

Rhodes, Russell. 2010. *A survey of vowel harmony with special focus on neutral segments, glides, and diphthongs*. Berkeley, CA: UC-Berkeley dissertation prospectus. http://linguistics.berkeley.edu/~russellrhodes/pdfs/Prospectus.pdf (accessed 07-May-2017).

Rubach, Jerzy. 1996. Nonsyllabic analysis of voice assimilation in Polish. *Linguistic Inquiry* 27. 69–110.

Schane, Sanford. 1984. The fundamentals of particle phonology. *Phonology Yearbook* 1. 129–155. DOI: https://doi.org/10.1017/S0952675700000324

Schwartz, Geoffrey. 2009. *Phonology for the listener and language learner*. Poznań: Wydawnictwo UAM.

Schwartz, Geoffrey. 2010. Phonology in the speech signal – unifying cue and prosodic licensing. *Poznań Studies in Contemporary Linguistics* 46(4). 499–518. DOI: https://doi.org/10.2478/v10010-010-0025-3

Schwartz, Geoffrey. 2013. A representational parameter for onsetless syllables. *Journal of Linguistics* 49(3). 613–646. DOI: https://doi.org/10.1017/S0022226712000436

Schwartz, Geoffrey. 2014. All gradience is not created equal. Ms. Adam Mickiewicz University. https://www.academia.edu/9320733/All_gradience_is_not_created_equal_-_chapter.

Schwartz, Geoffrey. 2016a. On the evolution of prosodic boundaries – parameter settings for Polish and English. *Lingua* 171. 37–74. DOI: https://doi.org/10.1016/j.lingua.2015.11.005

Schwartz, Geoffrey. 2016b. Representing non-neutralization in Polish *sandhi*-voicing. In Jolanta Szpyra-Kozłowska & Eugeniusz Cyran (eds.), *Phonology, its faces and interfaces*, 103–122. Bern: Peter Lang.

Schwartz, Geoffrey, Daria Arndt, Jerzy Dzierla & Ewelina Wojtkowiak. 2017. Pre-voicing variability in Polish and Polish English – evidence for a new Laryngeal Realism. *Paper presented at the 47th Poznań Linguistic Meeting*.

Schwartz, Geoffrey & Grzegorz Aperliński. 2014. The phonology of CV transitions. In Jolanta Szpyra-Kozłowska & Eugeniusz Cyran (eds.), *Crossing Phonetics-Phonology lines*, 277–298. Newcastle-upon-Tyne: Cambridge Scholars Publishing.

Smith, Jennifer. 2002. *Phonological augmentation in prominent positions*. Amherst, MA: University of Massachusetts at Amherst dissertation.

Steriade, Donca. 1993. Closure, release, and nasal contour. In Marie Huffman & Rena Krakow (eds.), *Nasals, nasalization, and the velum*, 401–470. San Diego, CA: Academic Press. DOI: https://doi.org/10.1016/B978-0-12-360380-7.50018-1

Stevens, Kenneth. 2002. Toward a model of lexical access based on acoustic landmarks and distinctive features. *Journal of the Acoustical Society of America* 111(4). 1872–1891. DOI: https://doi.org/10.1121/1.1458026

Stevens, Kenneth & Dennis Klatt. 1974. Role of formant transitions in the voiced-voiceless distinction for stops? *Journal of the Acoustical Society of America* 55(4). 653–659. DOI: https://doi.org/10.1121/1.1914578

Syrdal, Ann & H. Gopal. 1986. A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America* 79. 1086–1100. DOI: https://doi.org/10.1121/1.393381

Traunmüller, Hartmut. 1990. Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America* 88. 97–91. DOI: https://doi.org/10.1121/1.399849

Traunmüller, Hartmut. 1994. Conventional, biological, and environmental factors in speech communication: A modulation theory. *Phonetica* 51. 170–183. DOI: https://doi.org/10.1159/000261968

van Alphen, Petra & Roel Smits. 2004. Acoustic and perceptual analysis of the voicing distinction in Dutch initial plosive: The role of prevoicing. *Journal of Phonetics* 32. 455–491. DOI: https://doi.org/10.1016/j.wocn.2004.05.001

van der Hulst, Harry. 2015. The laryngeal class in RcvP and voice phenomena in Dutch. In Johanneke Caspers, Yija Chen, Willemijn Heeren, Jos Pacilly, Niels Schiller & Ellen van Zanten (eds.), *Above and beyond the segments*, 323–349. Amsterdam: John Benjamins Publishing Company.

Vaux, Bert & Bridget Samuels. 2005. Laryngeal markedness and aspiration. *Phonology* 22. 395–436. DOI: https://doi.org/10.1017/S0952675705000667

Walley, Amanda & Thomas Carrell. 1983. Onset spectra and formant transitions in the adult's and child's perception of place. *Journal of the Acoustical Society of America* 73. 1011–1022. DOI: https://doi.org/10.1121/1.389149

Wetzels, W. Leo & Joan Mascaró. 2001. The typology of voicing and devoicing. *Language* 77(2). 207–244. DOI: https://doi.org/10.1353/lan.2001.0123

Wright, Richard. 2001. Perceptual cues in contrast maintenance. In Elizabeth Hume & Keith Johnson (eds.), *The role of speech perception in phonology*, 251–277. San Diego, CA: Academic Press.

Zając, Magdalena. 2015. *Phonetic convergence in the speech of Polish learners of English*. Łódź: University of Łódź dissertation.