Acta Linguistica Academica 72 (2025) 1, 20–45
DOI: 10.1556/2062.2025.00852

A Neo-Jakobsonian merger of aperture, [−ATR], lowering, emphaticness, and retroflexion

MARKUS A. PÖCHTRAGER*

University of Vienna, Austria

Received: July 8, 2024 • Revised manuscript received: January 25, 2025 • Accepted: January 25, 2025

ABSTRACT

I argue that the system of Standard Yoruba vowels and their harmonic behaviour follow directly from the internal structure of vowels assumed in Government Phonology 2.0, a theory which reinterprets aperture (openness) as empty structure. I address a common connection between aperture and [−ATR] and propose that while aperture is empty structure in general, [−ATR] is a particular position within that empty structure. As a result, both Yoruba vowel harmony as well as other typical limitations of [−ATR] (e.g. in high vowels) become derivable. What is more, the same reasoning (and the same solution) can be applied to the representation of consonants, in languages as diverse as French, Swedish, and Arabic, allowing me to unify (at least) aperture, [−ATR], lowering, emphaticness, and retroflexion.

KEYWORDS

ATR, aperture, rhotics, retroflexes, emphatic consonants, structure

1. INTRODUCTION

This paper presents a rather ambitious enterprise: To show that aperture, [−ATR], lowering, emphaticness, and retroflexion (and possibly further properties) are all realisations of what is essentially (though not completely) the same underlying property, viz. empty structure. The unification of *some* of those properties has been argued for before, cf. Jakobson, Fant & Halle (1969) for emphatics and retroflexes (amongst others), and Schane (1990) for a link between aperture, [−ATR], and laxness. The present proposal takes this one step further (i) by combining

* Corresponding author. E-mail: markus.poechtrager@univie.ac.at

those two strands of research and (ii) by arguing that what those properties share is not a melodic property (as assumed by both Jakobson and Schane), but a structural one, in a sense to be made precise as the discussion unfolds, but in particular in sections 2.3–2.4.

The first claim, that of unification of several phonological categories, raises the obvious question how all those properties can still be distinguished (to the extent that they do need to be distinguished). For example, aperture (openness) and [ATR] play different roles in vowel inventories; the latter category might be involved in harmony to the exclusion of the former. I am going to argue that empty structure is what unites the two, but that there is a difference in the exact location of that structure within a phonological representation, which allows crucial distinctions to be upheld—we can have our cake and eat it, too.

Since a detailed and in-depth analysis of all the phenomena that form the basis of my argument to link the aforementioned properties would require more room than a single journal article affords, some parts will necessarily have to be brief. In section 2 I will start with openness and [ATR], both associated with vowels, where the connection is easiest to show. In particular, my demonstration will be based on ATR-harmony in Standard Yoruba. After going over the basic facts of harmony, discussing some previous analyses, and presenting the outlines of the theory to be employed here, I will show how and where aperture and [ATR] can be linked. From there, I will gradually extend the proposal to properties associated with consonants (lowering triggered by rhotics, retroflexes, and emphatics) in section 3.

The second claim, that the underlying nature of that all those properties is empty structure, follows from assumptions made in Government Phonology (GP) 2.0 (references in section 2.3), which this analysis is couched in, and which it extends. GP 2.0 argues that many properties commonly understood as melodic (i.e. by reference to elements in GP) should be reinterpreted in structural terms. This concerns not only the elements **ʔ** (occlusion) and **H** (voicelessness, aspiration, also high tone), whose replacement by structure marked the beginning of GP 2.0 (Pöchtrager 2006), but also the element **A** (openness, coronality). The element **A** displayed unusual properties (briefly revisited in section 2.3), many of which had to do with structure, which have led to the claim that **A** really *is* (empty) structure: As a consequence, any object that was thought to contain **A** must be reinterpreted as involving more empty structure, no matter what that element **A** encodes (openness, coronality, etc.). As shall become particularly clear in section 2.4, it is exactly this reliance on the relatively abstract notion of empty structure that allows us to unify various properties, while at the same time keeping them individually addressable.

Section 4 will contrast my account to previous work by Schane on the different realisations of the particle **a** (Schane 1990), which I follow to some extent, and by Jakobson on the feature [flat] (Jakobson, Fant & Halle 1969). I will argue that some shortcomings in those previous works can be successfully avoided in the account presented here. Lastly, section 5 touches on an open question in ATR-harmony and concludes.

2. YORUBA AND THE REPRESENTATION OF VOWELS

2.1. Yoruba vowels

Standard Yoruba (SY) has 7 oral vowels; the higher four are [+ATR], the lower 3 [−ATR] (Bamgbọṣe 1967; Archangeli & Pulleyblank 1989), as shown in (1). All Yoruba examples in this article will be given in regular orthography, only ⟨*ạ*⟩ with an additional diacritic deviates from (standard) ⟨*a*⟩. This is meant to facilitate visual recognition of [−ATR] vowels, all of which now have a dot underneath.

(1) SY oral vowels

i	*u*	
e	*o*	[+ATR]
ẹ ɛ	*ọ* ɔ	[− ATR]
ạ		

The mid-vowels stand in a special relationship with one another, as can be seen in the system of ATR-harmony active in the language. The relevant descriptive generalisation is summarised in (2), and then unpacked and illustrated in (3)–(5).

(2) a. The rightmost non-high vowel is unrestricted with respect to ATR.
b. A mid-vowel will be [−ATR] *iff* it is immediately followed by another [−ATR] vowel (mid or low).

(3) a. *ẹsẹ̀* 'foot' *ọbẹ̀* 'soup' *epo* 'oil' *olè* 'thief'
b. **esẹ̀* **obẹ̀* **ẹpo* **ọlè*

All words in (3a) end in mid-vowels, which can be either [−ATR] or [+ATR] as per (2a). Comparison between (3a) and (3b) reveals that a preceding mid vowel will pick up the same ATR-value, as stated in (2b).

The particular wording of (2b) also captures the fact that *ạ* acts as a trigger of harmony (4a), but not as a target (4b), i.e. it is unaffected by what is to its right:

(4) a. *ẹ̀pạ̀* 'groundnut' **epạ̀*
b. *ạ̀jẹ̀* 'paddle' *ạ̀te* 'hat'

Finally, (2) does not restrict the distribution of high vowels: They are neutral in that they precede (5a) and follow (5b) both [−ATR] or [+ATR] vowels. (5b) also explains the particular wording of (2a), since the rightmost non-high vowel ([−ATR] or [+ATR]) in each word is freely followed by a ([+ATR]) high vowel. High vowels are not only neutral but also opaque and block harmony; in (5c) between the final and the antepenultimate vowels:

(5) a. *ilé* 'house' *ilẹ̀* 'land' (neutral)
b. *ebi* 'hunger' *ẹ̀bi* 'guilt'
ạ̀rún 'five'
c. *Yorùbạ́* 'Yoruba' **Yọrùbạ́* (opaque)
odídẹ 'Grey Parrot' **ọdídẹ*

At this point, a number of (interconnected) questions have accumulated; amongst them: Why are all the higher vowels [+ATR] and all the lower ones [−ATR]? What is the connection between those two properties?[1] Related to that, what is the exact nature of the relationship holding between the two types of mid vowel? And finally, why are high vowels neutral? Within the GP literature, several proposals have been made addressing those questions and giving an account how to best analyse ATR-harmony in SY. I will look at these in the next subsection.

2.2. Previous analyses within GP

Ọla (1992) distinguishes an **A** element expressing aperture from a separate **ATR** element that corresponds to [+ATR] in binary feature systems. ([−ATR] is then the absence of that element.) The gist of the analysis is that an **ATR** element can only spread if both target and trigger *also* contain the element **A**.[2] This explains why mid-vowels (which contain **A**) will be sensitive to the following vowel, as seen above in (3)–(4): If another mid-vowel (which of course also contains **A**) follows, an **ATR** element will spread from that following vowel to the left. If a low vowel follows (again containing **A**), the mid-vowel will surface in its (underlying) [−ATR] shape, as low vowels do not contain an **ATR** element, so there is nothing to spread.[3] Lastly, it also explains why high vowels, while containing the **ATR** element, are neutral: They do not contain **A**, so the conditions for spreading will never be met.

Ọla's analysis captures the data, but raises the obvious question *why* spreading of **ATR** should be contingent on the presence of **A** in both target and trigger? Ọla assumes that **A** elements in subsequent nuclei undergo the Obligatory Contour Principle (OCP) and that that is what triggers spreading of **ATR**. This changes the question slightly, but does not answer it. Why should an OCP effect on one element have consequences for a completely unrelated element?

Cobb (1997, 2003) assumes that there is no **ATR** element per se in the universal set of elements. Instead, ATR is encoded by headedness: If any one element within the vowel takes over the role of head, the vowel is [+ATR], otherwise [−ATR]. Headedness can be passed on from right to left (so-called "head alignment"), on the condition that the trigger is complex, i.e. contains more than one element.[4] This immediately excludes the low vowel and the high vowels as

[1]The system in (1) has provided a fertile ground for analyses involving underspecification (Archangeli & Pulleyblank 1989), since [ATR] is only contrastive for mid-vowels. Non-mid vowels are underspecified for ATR in such analyses, yet at least [−ATR] needs to be filled in in time to make ạ a trigger, casting doubt on the usefulness/testability of underspecification (Dresher 2009, 125f). No underspecification is needed (or indeed possible) in the account to be developed in this article.

[2]Charette & Göksel (1996, 49f) discuss a similar case in the Turkic language Sakha, where spreading of **U** is contingent on the presence of **A** in both target and trigger. They refer to this as an "**A**-bridge".

[3]Ọla (1992) worked with an old version of element theory where each element had a certain charm value, which was supposed to account for why certain elements combine more easily than others. Both **A** and **ATR** had positive charm, and would thus repel each other. From this it follows that ⟨ạ⟩, where **A** is the only element and the head, thereby determining the charm value of the entire vowel, will not combine with **ATR**, neither lexically nor as the result of spreading. Mid-vowels also contain **A**, but not in head position, which is why the addition of **ATR** is licit.

[4]Cobb (2003) attempts to derive the condition that triggers must be complex from the Complexity Condition, according to which a governor has to be more complex (Kaye, Lowenstamm & Vergnaud 1990, 218) or at least as complex (Harris 1990, 274) as a governee. Head alignment is seen as a kind of government, and complexity simply refers to the number of elements contained in a particular segment. However, the Complexity Condition is usually seen as relative, i.e. comparing the complexity of two segments, instead of absolute, i.e. requiring that a governor be complex or contain a specific number of elements.

triggers, as each one of them contains a single element only; only mid-vowels will be able to pass on their headedness to the left.[5] As for the targets, high vowels will be unaffected as they are lexically headed, while the low vowel is exempt by the the stipulation that **A** (the only element it contains) cannot be made head. That last stipulation raises a similar concern as with Ọla's account: Why should **A** be unwilling to be head (with **A** failing to be head the vowel will be [−ATR])? Why and how are those two properties connected? Conversely, what is it about the high vowels that they are lexically headed (i.e. [+ATR])? Similar questions could be raised about Backley (1997), who employs a variant of the headedness account.

What all those accounts have in common is the assumption that openness (mediated by **A**) is separate from ATR (however expressed), yet those two properties interact in all analyses. At this point it is as well to remember the non-negotiable core assumption of GP, viz. the Non-Arbitrariness Principle, which demands a connection between a phonological process and its context. As long as openness and ATR are treated as properties that are independent of one another, such a connection will be lacking.

One GP account that manages to avoid this particular problem is that of Polgárdi (1998, 141ff). In her quite interesting analysis, the three lower vowels (the [−ATR] set in (1)) contain **A**, while the four higher vowels (the [+ATR] set) lack it. ATR does not figure in Polgárdi's analysis at all; hence the question why **A** would go together with ATR does not arise. The four higher vowels (all lacking **A**) are internally distinguished by whether **I**/**U** is the head (*i*/*u*) or not (*e*/*o*), and harmony is then simply the spreading of **A**, which fails to target (*i*/*u*) due to Structure Preservation (spreading of **A** onto high vowels would give rise to a combination of elements that does not occur lexically). Polgárdi's complete replacement of ATR by **A** has certain similarities to my proposal in section 2.4, but there are two crucial differences: Firstly, in my account ATR can still be properly singled out, while at the same time ATR and **A** *are* made very much identical in many other respects. Secondly, Polgárdi's solution has no (obvious) application beyond (SY) vowel harmony, while the proposal in section 2.4 does go well beyond vowels or harmony, since it involves a more wholesale replacement of **A**.

2.3. The element formerly known as A

Within the wider GP literature there is some discussion what the element **A** should represent in consonants. Some (amongst others: Broadbent 1991; Cyran 1997; Goh 1997) argue that **A** should encode coronality, a view that does not meet with unequivocal approval: Backley (2011) argues for **I** or some combination of both **A** and **I**; those arguments are addressed (and contested) in Pöchtrager (2013). For the purposes of this paper I will assume that the identification with coronality is essentially correct; and the analyses in section 3 point in that same direction.

Extensive arguments that **A**, both in vowels (openness) and in consonants (coronality), should be replaced by structure have been presented elsewhere, notably Pöchtrager (2006, 2010, 2012, 2013, 2016, 2018a, 2020, 2021a,b, 2023, 2024). The gist of the argument is as follows: **A** in general behaves quite differently from other elements; more precisely, it seems to interact

[5]This predicts that a form like the one marked as ungrammatical in (4a) should actually exist. Cobb (2003) gives some such forms that she has elicited from a speaker, but without further discussion of their status. This is in contradiction to other sources, e.g. Archangeli & Pulleyblank (1989).

with constituent structure. This interaction can be characterised such that **A** often provides extra room; it allows bigger constituents than otherwise expected. Such interaction with structure suggests that **A** is structural itself. This is easily illustrated with size restrictions in English monosyllables: Ignoring any initial onsets, we either have a long vowel/diphthong and a single consonant (*meet*, *boot*, *boat*), or a short vowel and a cluster (*mint*, *lift*, *pact*). The limit can be exceeded, though, if both consonants are coronal (i.e. contain **A**), in which case a long vowel/diphthong can be followed by a cluster: *fiend* (but no **fiemp*, **fienk*), *count* (but no **coump*, **counk*), *east* (but no **easp*, **eask*), etc.[6] This general pattern has various subpatterns, one of them being the following: Southern British English allows **ɑ:** (i.e. **A** by itself) before a cluster of fricative and consonant even if only *one* of the consonants is coronal (contains **A**): *clasp*, *task*, *draft*. Other vowels do not have this freedom: **cleesp*, **toosk*, **dreeft*. This is revealing inasmuch as the vowel (containing **A** and nothing else) seems to be able to make up for the "insufficiency" of the cluster (which only contains one coronal), as long as there are two segments with **A** around (the vowel itself and one consonant). What is more, this additional freedom given by **A** is not restricted to English, but recurrent across languages (Pöchtrager 2012): Finnish has long vowel plus coronal cluster as in *aalto* 'wave' (but no **aalpo*, **aalko*), Spanish (Harris 1983) has *veinte* '20', *treinta* '30', or *fausto* 'pomp' with a diphthong before a coronal cluster, etc.

This is not a novel observation or problem, but the solutions proposed in the literature usually amount to providing special syllabic positions for coronals (Fudge 1969; Selkirk 1982; Vaux & Wolfe 2009). Obviously, simply moving coronals to dedicated positions does not explain *why* coronals are special to begin with. GP 2.0 tries to avoid that by integrating this special property right into the representation. **A** is reinterpreted as empty structure: Objects that were thought to contain **A** before (non-high vowels, coronals) will instead contain more (empty) positions than those without. Those empty positions can in turn explain why **A** seemed to be so conducive to superheavy structures: In a word like *fiend* the vowel borrows space from the coronals, since those are characterised by several empty, unused positions that are available to adjacent segments. This makes a long vowel possible where none should occur (Pöchtrager 2010).[7]

Additionally, such a reinterpretation of **A** also yields a scalar interpretation of vowel aperture: high vowels are relatively small, close-mid vowels bigger, and open-mid vowels even bigger. In several varieties of Brazilian Portuguese stressed **ɛ** reduces to **e** when unstressed, and further to **i** in final unstressed position. This can be modelled as the successive removal of layers of structure in progressively more unfavourable prosodic positions (Pöchtrager 2018a).

After this very general recapitulation of the reasons for rejecting **A** as an element we can look at the formalism defining the shape of phonological representation in GP 2.0. As in earlier versions of GP (e.g. in Kaye, Lowenstamm & Vergnaud 1990), skeletal positions function as the mediator between melody and structure, as skeletal positions (i) can be associated with melodic material and (ii) are grouped into constituents. There are some important differences, however. For example,

[6] Words like *fiend* are sometimes said to contain a "superheavy rhyme"; but since the sequence **i:nd** does not form a constituent in (any version of) GP, "superheavy structure" would be more adequate.

[7] The formal mechanism underlying this is melodic command (m-command), which fulfils a role similar to spreading in more traditional autosegmental theories. That is, a long segment involves a head that m-commands (an) empty position(s), as a result of which all involved positions contribute to the length of the segment. This is also how length before coronals is captured. M-command will play no further role in this contribution; details, also about the difference to spreading, can be found in Pöchtrager (2006) and Kaye & Pöchtrager (2013).

since in GP 2.0. the focus is shifted from melody to structure, with structure playing a more important role, the number of elements has shrunk consistently. Even more importantly, what used to take up a single skeletal position in the past (e.g. the consonant at the beginning of *tea*, represented as a single skeletal position dominated by a non-branching onset) will now be considerably more complex in terms of skeletal positions and their organisation. This point in particular will become more obvious once we get to the representation of vowels, which typically consist of several skeletal positions. Since I cannot go into a justification for each point of difference between older and newer theory here, the reader is referred to Pöchtrager (2006, 2021b) for detailed discussion, as well as Kaye & Pöchtrager (2013) for a contrastive view between different versions of the theory. (6) lists some general axioms of the theory that will be relevant for this article, but does of course not represent a complete axiomatisation of the theory.

(6) a. Skeletal points come in two kinds, heads or non-heads.
b. Heads subdivide into two groups with two members each: Two types of nuclear heads (xn, xN) and two types of onset (non-nuclear) heads (xo, xO).
c. Skeletal positions can be annotated with elements, of which there are three left in the theory: **L** (nasality/true voicing/low tone), **U** (labiality), **I** (palatality).
d. A skeletal position can be annotated with maximally one element.

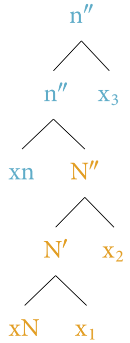
Since the first part of this contribution deals with the internal structure of vowels, we will need to look at the role of nuclear heads in more detail. (7) gives further assumptions that pertain to the internal structure of nuclei.

(7) a. A nucleus consists of maximally two heads (xn, xN).
b. xn can select a non-head x or (a projection of) xN;
xN cannot select (a projection of) xn.
c. Each head can project up to twice (xn/n′/n″, xN/N′/N″).

Nuclei in GP 2.0 are taken to have a bipartite structure, i.e. they can be made up of two nuclear heads, as per (7a). Those nuclear heads are slightly different in what exactly they contribute to a phonological representation[8] and also with respect to their selectional properties: xN (or a projection thereof) can be embedded inside a projection of xn, but not the other way round. There is no requirement that both heads are present: Either one could be missing as long as there is at least one head to make sure that the structure is identifiable as a nucleus. But if both are present, (the projection of) xN will be embedded in that of xn. (I will occasionally refer to xn as the higher head and xN the lower head, even if only one is present.) All other positions in the tree (marked ‘x’) are non-head positions. All this, together with (7c), inspired by the x-bar schema, yields the maximal structure in (8). The colours are a visual aid and have no significance beyond that. Likewise, the indices on non-head positions only help identification but have no theoretical importance.

[8]The difference is not immediately relevant here. Pöchtrager (2023) argues that xn is the formal representation of metrical prominence; for a language like English: stress. Note that a bipartite structure for the nucleus with two heads, but of a very different kind, is also assumed in den Dikken & van der Hulst (2020).

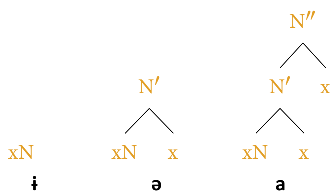
(8) Maximal structure of a nucleus



This is a fairly complex structure consisting of five skeletal positions, of which two are heads (xn, xN) and three non-heads (x_1, x_2, x_3). The notion of projection expresses constituency, such that xN is closest to x_1, etc. It also says that an entire constituent (say, N′) is of the same type as its head xN, just like in syntax. The vowels of individual languages are represented by various substructures of (8), as exemplified in (9)–(10). The *amount of empty positions* in a nucleus encodes openness, which is what (9) focuses on. Each of the three structures counts as a nucleus, thanks to the head (and projection).

With this, the old element **A** becomes superfluous as an expression of openness.[9] The maximal structure in (8), as a corollary of the assumptions in (7), also makes predictions about how many degrees of aperture can be distinguished in a language. The way the theory achieves this is by limiting the number of heads and their selectional requirements, and the number of times each head can project. Given that no there is an upper limit on the number of degrees of height that languages distinguish, such an inbuilt cut-off point is necessary.[10]

(9)

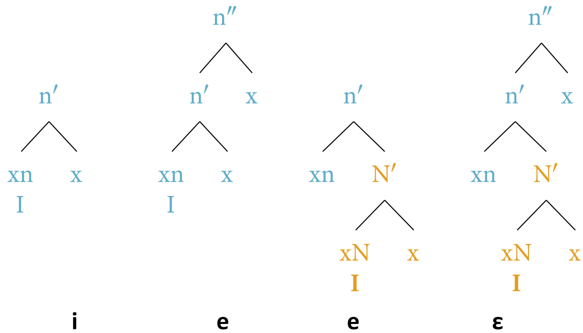


Lastly, as (6c) states, skeletal positions can be annotated with elements. In the context of this article I will focus on heads annotated with elements; non-heads annotated with elements will mostly (but not exclusively) be relevant for diphthongs, cf. Živanović & Pöchtrager (2010), but also Pöchtrager (2024). (10) gives some examples of front vowels, all of which contain **I**.

[9]In fact, the structures of GP 2.0 do more than just replace the old element **A**. For one thing, they allow a more fine-grained, scalar representation of aperture. Also, they cover more than what old **A** would have covered, as this article (as well as Pöchtrager 2023) is meant to show.

[10]Not only do the assumptions in (7) restrict the number of possibilities, they also restrict them to what seems to be the correct number: Danish (Pöchtrager 2018a) has a relatively rich inventory of vowels, with four degrees of (long) front unrounded vowels, typically given as **iː/eː/ɛː/æː** (Basbøll & Wagner, 1985, 50). This is exactly what (8) allows as a maximum. Similar remarks can be made about the inventories listed in typological work, e.g. Crothers (1978), though of course each system would need to be looked at separately.

(10)

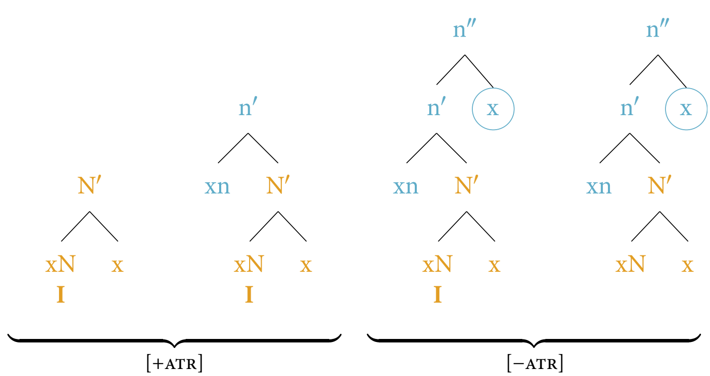


This again illustrates how empty position are used for the expression of aperture. The high vowel **i** contains one empty position, close-mid **e** two, and open-mid **ɛ** three.[11] It bears repeating that it is the number of empty positions that matters for the expression of openness, not simply the size of the tree: Both **ɨ** in (9) and **i** in (10) contain one empty position each and thus count as high. Note also that the two representations for **e** in (10) are no accident: The theory allows for several possibilities for what might be considered one and the same sound. Such structural ambiguities allow the modelling of (often cross-linguistic) behavioural differences, cf. Pöchtrager (2021a, 2023) for discussion of this point in particular.

2.4. The first unification

With the theoretical preliminaries out of the way, I will now proceed to an analysis of SY vowels and vowel harmony. (11) gives the GP 2.0 representations of SY *i*, *e*, *ẹ*, *ạ*, as these will be sufficient for the discussion. Again, the colours are simply a visual aid; the circled position will be explained below.

(11) a. *i* b. *e* c. *ẹ* d. *ạ*



[11] One might wonder why high vowels like **i** contain an empty position at all, instead of no empty position, as there are no vowels that are even higher. That position is necessary to express tenseness/laxness, as argued in Pöchtrager (2020).

Openness is directly captured by the number of empty positions. (Note how *ẹ* and *ạ* are equal in size but unequal in the number of empty positions.) The vowels *u*, *o*, *ọ* will be parallel to (11a–c), but with **U** instead of **I**.[12]

The representations in (11) follow a certain inner logic, i.e. the set in (11) is not just some random collection of structures. The assumptions determining their composition are given in (12), and I hasten to add that those are independent of (and therefore not crucial to) the analysis of harmony. At this point they are pure stipulations that are intended to limit what is universally possible to what actually occurs in SY. (They are similar in kind to the Licensing Constraints of earlier versions of GP, cf. e.g. Charette & Göksel (1996).) Their stipulatory nature is probably most obvious in the case of (12c), whose sole purpose is to exclude central vowels other than *ạ*. I will also leave open whether and to what extent the subclauses are interconnected.

(12) a. Only xN can be annotated with melody.
b. xN projects exactly once.
c. If xN is not annotated with melody, xn projects up to n″.

Those three assumptions circumscribe the shapes of the trees we find in SY. Note that the theory (in general) only allows one element per position, cf. (6d). This, in conjunction with (12a), excludes front rounded vowels as they would require both **I** and **U**, but only the lower head xN can be annotated with elements.[13]

Now, what is more important in the context of this article, the particular representations chosen in (11) also allow for a formal identification of ATR, as expressed in (13).

(13) The specifier of xn (Specxn) is the formal representation of [−ATR].

[12]The representation of *ạ* in (11d) is slightly different from that of **a** in (9). This is a property that GP 2.0 shares with Particle Phonology, where the exact representation of the most open vowel in a language will be slightly different depending on how many degrees of aperture the other vowels in the system display. Schane (1984, 139f) refers to this as the "the law of maximum aperture". In GP 2.0, it is generally the case that considerations of phonological behaviour and distribution play a role in determining the exact representation of individual segments; cf. the discussion of (10). Earlier versions of GP also had a similar kind of leeway when it came to headedness: In a combination of two elements *x* and *y*, either *x* or *y* or neither could be head, and the final decision would typically depend on behaviour, including e.g. membership in natural classes. In fact, SY does not *enforce* a representation of *ạ* as in (11d); that vowel could simply lack N′ and have a non-head x as the sister of xn. This would leave the analysis of harmony (below) unchanged; it would only have consequences for the assumptions in (12) and imply that any differences in openness between *ẹ* and *ạ* are phonologically irrelevant.

[13]Languages with front rounded vowels will have those two elements in different positions, cf. Pöchtrager (2017, submitted) for a discussion of Finnish **y**/**ø**, with **U** in the lower head xN and **I** in the higher head xn.

The specifiers of xn are the positions circled in (11). This divides the set in two halves: The vowels in (11c–d) contain a Specxn and thus count as [−ATR]; those in (11a–b) do not and are therefore [+ATR].[14]

In this system, [−ATR] is structurally more complex than [+ATR], in that all and only [−ATR] vowels will contain an empty specifier of the higher head xn. This allows us to unify the old element **A** and ATR: Both are expressed by more empty structure. At the same time we can still single out ATR, as it is expressed by a *particular* empty position in the tree, viz. the specifier of xn. With those representations in place, SY vowel harmony can then be expressed as follows:

(14) If a vowel projects up to Specxn (= [−ATR] *ẹ*, *ọ*, *ạ*), the vowel to its left will do as well, provided it contains xn to begin with.

This is somewhat similar in spirit to the head-alignment accounts discussed in section 2.2, in that subsequent vowels will agree; not with respect to headedness, but with respect to whether they contain an (empty) Specxn. The details of harmony follow from the representations in (11) together with the statement in (14). Let me go through the three relevant cases individually.

Firstly, a mid vowel preceding the three triggers (*ẹ*, *ọ*, *ạ*) must contain an empty Specxn, i.e. it will be [−ATR] *ẹ*, *ọ*. Put differently, if *e* or *o* come to stand in front of one of the triggers, Specxn of the trigger will be copied over to the left onto the target, deriving [−ATR] *ẹ*, *ọ*.

Secondly, if *ạ* precedes one of the three triggers, nothing will happen, as *ạ* already contains a Specxn and thus meets the requirement vacuously.

Thirdly, high vowels do not contain a higher head xn and thus stay unaffected when followed by one of the triggers. The rationale behind this is that there cannot be a specifier to a particular head if that particular head is not part of the representation to begin with; in other words, Specxn could not be copied over to the left if there is no xn in the target vowel. This explains both why high vowels will not harmonise and why they block the propagation of harmony (assuming that harmony cannot skip vowels).

One detail remains to be looked at. (5c) showed that high vowels are opaque and block the spreading of [−ATR]: *Yorùbạ́* 'Yoruba', **Yọrùbạ́*. The account does not yet explain why **Yọrùbạ́* (or any sequence [−ATR]–high vowel–[−ATR]) should be ungrammatical; after all, the first vowel could be lexically [−ATR]. The only way to exclude this is to make an assumption similar to the one in Archangeli & Pulleyblank (1989) (and also in Polgárdi (1998)), viz. that [−ATR] (in the present account: Specxn) is a morpheme-level property that is associated with the right-most potential host and is passed on through the domain from there.

Before moving on to other properties of ATR-systems that can be derived from the proposal in (13), I would like to highlight some general aspects of the analysis:

[14] A reviewer queries why the presence of Specxn is taken as the expression of [−ATR], instead of n″, since the two always go together. The answer is that Specxn is a skeletal position, while n″ is simply a constituent label. As argued in the previous section, empty skeletal positions are an expression of openness, whether that be openness as understood when we talk about vertical tongue position (leading to higher F_1), or openness as brought about by [−ATR] (also leading to higher F_1), cf. Lindau (1978). The difference is only in the location of the empty position within the tree. The fact that an exchange between the two types of openness is possible (section 2.5) supports this view, as such an exchange is easily expressible by simply having an empty skeletal slot integrated in a different position.

Firstly, having a specifier (trivially) requires having a particular head, i.e. the correct “foundation” is needed to build upon; vowels without xn cannot have Specxn. I will come back to this concept in section 3.2.

Secondly, the bipartite structure of nuclei, and in particular the existence of two heads, is crucial for the analysis: A vowel like SY *i* in (11a) does have *some* nuclear head (after all, this is what guarantees that it will be a nucleus), but that happens to be the lower head xN. It lacks the higher head xn whose specifier would represent [−ATR]; thus I derive why certain vowels are neutral/opaque.

Thirdly, the bigger a structure is (i.e. the more open a vowel is), the more likely is it that there will be an upper head xn (and a specifier to that head), for the simple reason that there is an upper limit on how many times a head can project, as given in (7c), and any particular vowel will extend over a substructure of that theoretical maximum. If a vowel is very open, then the empty positions contained in the projection of a single head, be it xn or xN, will not be enough; the vowel will necessarily extend over two heads. This predicts that more open vowels are more likely to be [−ATR]. And inversely, the fact that the high vowels of SY are all [+ATR] is directly connected to their (small) size.[15]

Fourthly, SY vowel harmony is expressed as a condition on specifiers having to be in agreement; it is via those positions high up in the tree that the vowels of a word “communicate” with one another. This is similar to an idea expressed in Rennison (1990) to the extent that the internal arrangement of elements in a segment determines their visibility to other segments in a domain. The same idea is also employed in the analysis of Finnish vowel harmony (Pöchtrager 2017, submitted), where both front and (harmonically) neutral vowels contain **I**, but an **I** contained in front vowels is always visible (it is high up in the tree), but in neutral vowels only under very specific conditions (when there are no other “big” vowels around, i.e. when neutral vowels are all by themselves, in which case they behave as if they were front). This asymmetry can be expressed quite naturally by exploiting the various positions in a tree that can be annotated with elements.

2.5. Further properties of ATR-systems

In some languages ATR plays an even more pervasive role than in SY. For example, Kalenjin (Lodge 1995) has five [+ATR] and five [−ATR] vowels.[16] This, in comparison to SY and its seven vowel system, highlights a noteworthy property of ATR-systems: While [−ATR] high vowels and [+ATR] low vowels can be found in principle, some languages avoid them. Put in more traditional phonemic terms (which I do not wish to endorse), ATR contrasts both in the low and the high vowels are somewhat unstable.

As for a missing [+ATR] low vowel, this is what Kaye (1980) referred to as the “mystery of the 10th vowel”: The [+ATR] counterpart of [−ATR] **a** can have various realisations (**æ∼ɐ∼ə∼ʌ∼ɛ**),

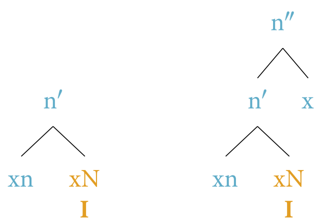
[15]This does not mean that the system completely excludes high [−ATR] vowels, only that they are somewhat unstable. I return to this in section 2.5.

[16]Lodge (1995) gives the [+ATR] system as **i e a o u** and simply uses an italic font for their [−ATR] counterparts. He also lists a whole series of phonetic correlates of ATR, which go well beyond the vowel proper and even involve phonation and manner characteristics of neighbouring consonants. While doubtlessly interesting in view of the aims of this article, viz. the unification of several phonetic categories, those properties will have to be left for further research.

including open-mid/mid vowels, which in turn might lead to a merger with the [+ATR] counterpart of another (mid) vowel. (Such a merger is what Kaye deals with.) Furthermore, this vowel is typically lost first.[17] Similarly, [−ATR] high vowels are diachronically unstable and tend to be lost. As already pointed out by Schane (1990), what is particularly interesting is that [−ATR] *high* vowels typically merge with [+ATR] *mid* vowels (Stewart 1971): **ɪ** > **e**, **ʊ** > **o**. (I will look at Schane's account more closely in section 4.)

The GP 2.0 proposal does offer a way to deal with those asymmetries. Since the second case (lowering of high vowels) is more straightforward than the first one (10th vowel), let me begin with the second one. Under the same assumptions that I have made so far in this article, viz. (i) openness is linked to empty structure, (ii) [−ATR] is expressed by Specxn, and (iii) melody (**I**) can only sit in xN (the latter being a language-specific stipulation, the same as in SY), there is really only one way to represent high (front) vowels distinguished by ATR, and that is as in (15):

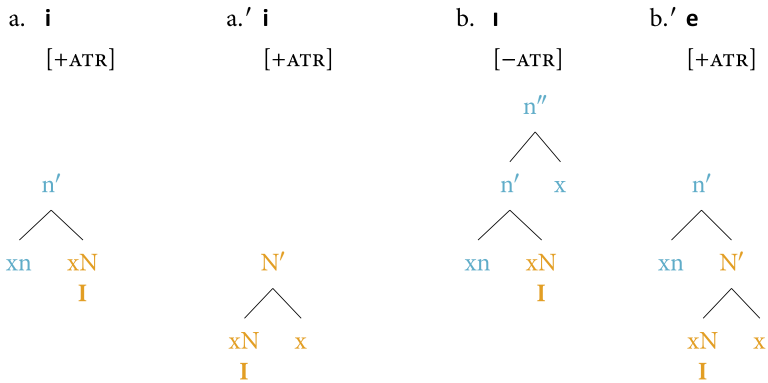
(15) a. **i** [+ATR] b. **ɪ** [−ATR]



(15a) has one empty position and will therefore count as high; (15b) is like (15a) plus a specifier of xn to make sure that the vowel is [−ATR]. Having a Specxn requires the appropriate head (xn); this is what section 2.4 referred to as the foundation. The lower head xN is required in order to host the melodic element; **I** in this case. As a result of all these considerations we arrive at the rather unusual structures in (15a–b), where we have two heads as sisters, since xN does not project. Assume now that this is a somewhat marked configuration that tends to be reinterpreted by learners. In that process (i) the total amount of structure is kept constant and (ii) xN is assumed to project once. The reasoning behind those two conditions is as follows: By keeping the amount of structure constant (though not necessarily the kinds of positions involved) the openness of the vowel will be minimally affected, if indeed at all. (That qualification will become clear anon.) Transmission across generations is for the most part relatively accurate, though reinterpretations can of course arise. Such a reinterpretation is what happens with regard to the question of which head can project. By assuming that xN projects once, the marked structure of two heads as sisters will be avoided. This scenario is illustrated in (16): (16a–b) repeat (15a–b), while (16a′) and (16b′) show what (16a) and (16b), respectively, will turn into.

[17] Akan is sometimes given as a 5 + 5 system like Kalenjin; but note that Clements (1985) argues that the [+ATR] low vowel in Akan is not phonemic.

(16)



The representation in (16a) involves one level of projection. If, in diachronic change, the marked structure is to be replaced by one without two heads as sisters, while keeping the amount of structure constant (two skeletal positions), then the only way to achieve this is by having the lower head xN project. As a result, (16a′) arises, i.e. a high vowel that is identical to SY **i**. Note that since both (16a) and (16a′) lack a Specxn, they are both [+ATR]. The change from (16b) to (16b′) is entirely parallel to the one from (16a) to (16a′) and has the same motivation: Sisterhood between two heads is avoided by having the lower head xN project; the total amount of structure (three skeletal positions here) remains the same. But the change from (16b) to (16b′) has more tangible effects: In (16b) xn projects twice, while xN does not project. If the goal is to reinterpret xN as projecting (once), while keeping the total amount of structure constant, Specxn will be lost and with it [−ATR]. In other words, the resulting structure in (16b′) is identical to the one assumed for SY [+ATR] **e**.[18] This is exactly the merger that we want: A former [−ATR] **ɪ** is reorganised in such a way that [−ATR] is lost but openness gained.

Let us now move on to Kaye's "mystery of the 10th vowel" mentioned above, which the GP 2.0 approach also allows us to make some degree of sense of. That "mysterious" 10th vowel is the [+ATR] counterpart of [−ATR] **a**, which has several puzzling properties. For one thing, that vowel is typically lost first (Stewart 1971), i.e. even before the [−ATR] high vowels, changing a 10- into a 9-vowel system. I have nothing to offer as an explanation for that particular order. In addition, the exact quality of that 10th vowel ranges over a number of possibilities (**ɐ/ə/ʌ/æ/ɛ**),[19] i.e. (open-)mid and sometimes also front. The fronting seen in **æ/ɛ** is hard to explain under current assumptions within (potentially all versions of) GP. Frontness is captured by the element **I**, so if the [+ATR] counterpart of [−ATR] **a** is something like front **æ**, do we need to posit an

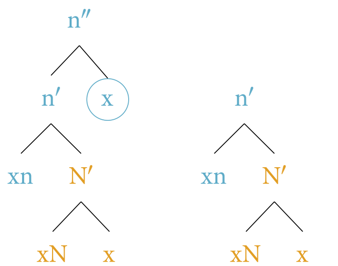
[18]And also the structure of **e** in a 5 + 5 system. Structure (16b′) also makes clear why non-high vowels like **e** are under no pressure to reorganise: The two heads are not sisters.

[19]I am aware that the symbol **ɛ** is often used to denote a [−ATR] vowel, yet here we are talking about the [+ATR] counterpart of [−ATR] **a**. There is no contradiction here; the vagueness/apparent confusion is a direct result of the set of symbols provided by the IPA; **ɛ** is not inherently [−ATR] (or "lax", for that matter), but simply a cover for anything close enough to cardinal vowel 3.

element **I**? Other ATR pairs (**o/ɔ**, say) do not involve an extra **I**; why should the low vowel do so?[20]

There is, however, one aspect of the mystery of the 10th vowel that follows quite straightforwardly from my account so far, and that is the *raising* involved in qualities like **ɐ/ə/ʌ**.

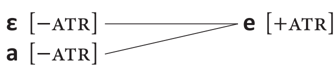
(17) a. **a** [−ATR] b. **ɐ** [+ATR]



(17a) gives the structure of **a**, which is identical to SY *ạ* in (11d). The vowel is [−ATR] as it contains a Specxn. It is a central vowel since it is not annotated with any elements. Its [+ATR] counterpart in (17b) must then lack that specifier, and given the fewer number of empty positions it will be less open, i.e. a kind of open schwa.[21]

I will finish off this section with the discussion of a case where the [+ATR] counterpart of [−ATR] **a** is fronted; but, as mentioned before, there is the problem of explaining where the element **I** comes from. This problem will remain unresolved, but the concomitant raising that is involved can be made sense of. Kaye (1980) discusses the Kru language Dida, which has five [−ATR] vowels, but only four [+ATR] vowels. As (18) shows, the [+ATR] counterpart of **ɛ** and **a** (both [−ATR]) is identical. (19) provides representations for some Dida vowels.

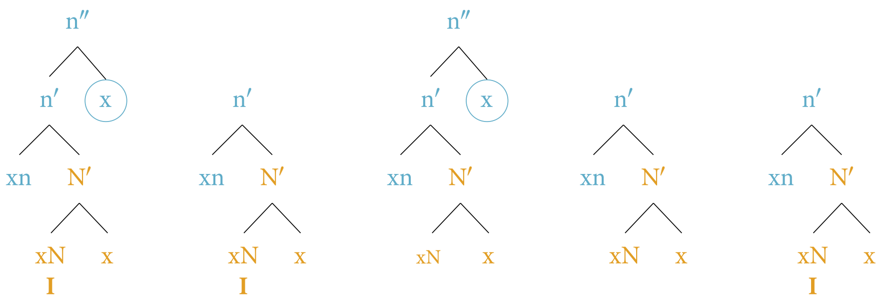
(18) Dida vowel counterparts following Kaye (1980)



[20] One might argue that there is no extra **I** and that the front quality is simply "a matter of interpretation". I would consider this a somewhat dangerous move, as it allows us to explain away any unexpected realisation of a phonological object. But it is true that there are other mysterious appearances of **I** with no obvious source, as in the palatalisation of Portuguese sibilants (**s** → **ʃ**) in non-prevocalic position. This suggests to me that our current theories are simply not sophisticated enough to understand what **I** really is and why it sometimes seems to appear out of nothing.

[21] With one empty position fewer it would end up as **ə**; such a representation for schwa is also supported by vowel reduction in various languages such as Portuguese or Catalan; cf. Pöchtrager (2018a) for details.

(19) a. **ɛ** b. **e** c. **a** d. illicit e. **e**



(19a–c) are identical in representation to SY *ẹ*/*e*/*ạ*. The vowels in (19a, c) are both [−ATR] due to Specxn. The [+ATR] counterpart of [−ATR] **a** in (19c) will lack that specifier position, i.e. we expect a representation as in (19d), identical to (17b). Assume now that (19d) is an ungrammatical structure in Dida, in the sense that the language has no central vowel that is less open than **a**, as it would necessarily be, given the fewer number of empty positions. In order to rescue that structure an **I** is introduced in the lower head. This guarantees the front quality, and at the same time raises the vowel even further, to mid, since the introduction of an element decreases the number of empty positions. The structure so remedied, (19e), is of course identical to (19b), i.e. the merger is captured.

There are two aspects to this mini-account that for the time being remain highly unsatisfactory (to me). Firstly, I seem to be saying that (19d) is ungrammatical because it is absent from the system, i.e. ungrammatical because it is ungrammatical, which is hardly insightful. Secondly, the appearance of **I** remains unmotivated, as discussed before. However, there is one very positive aspect, and that is the connection between the appearance of an element and the (additional) raising of the vowel. This connection falls out straightforwardly from the setup of the theory: Introduction of an element decreases the number of empty positions, making the vowel less open.

3. THE CONSEQUENCES FOR CONSONANTS

3.1. The element A in consonants

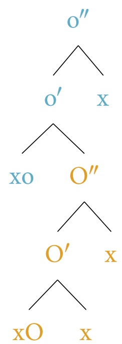
As discussed in section 2.3, there is some debate in the literature whether the element **A** should really represent coronality. But note that no matter whether **A** stands for coronality or something else; if it stand for anything at all, then its abolishment will have repercussions for the representation of that property as well. Based on the arguments in section 2.3 I will assume that the identification with coronality is essentially correct; and the phenonema to be discussed below do in fact lend further support (albeit indirectly) to that identification. This means that coronality, like openness, must be characterised by more empty structure; cf. Pöchtrager (2021b) for detailed discussion of that same point.

Onsets in GP 2.0 have a bipartite structure entirely parallel to nuclei;[22] (20) differs from (7) only by type of heads involved.

(20) a. An onset consists of maximally two heads (xo, xO).
b. xo can select a non-head x or (a projection of) xO;
xO cannot select (a projection of) xo.
c. Each head can project up to twice (xo/o′/o″, xO/O′/O″).

As with nuclei, this gives the maximal structure in (21).

(21) Maximal structure of onsets



Place of articulation can the be expressed as follows: Both coronals and velars will be devoid of elements, unless as a result of secondary articulation. (Note that what is said about velars here might extend to uvulars and glottals.) But coronals will involve bigger structures, while velars are smaller structures.[23] In other words, this parallels the distinction between low vowels and high vowels: What sets them apart is the amount of (empty) structure. Now, “bigger” and “smaller” are somewhat vague and need clarification. In Pöchtrager (2021b) it was assumed that the lower head xO explicitly expresses coronality, i.e. a structure with xO would automatically be coronal, one without would not. There is reason to believe, coming from a parallel problem in vowels (Pöchtrager 2021a, 2023), that this identification of xO as coronal is problematic and should be rejected. Rather, a consonant with both xo *and* xO is coronal; a consonant with only one of the two heads (no matter which) is not. Last but not least, labials are also smaller structures, like velars, but unlike them they also contain the (labial) element **U**. This allows for various groupings on formal grounds that correctly model common patterns: Firstly, coronals *and* velars are both devoid of melody; they are structure only, but differ in size. The fact that they are both considered as relatively unmarked (Paradis & Prunet 1991) can be attributed to the lack of melody. Secondly, velars and labials are united by their relatively small size (in contrast to

[22] “Onset” as employed here (or in GP in general) does not necessarily overlap with what is called an onset in mainstream theories of syllabic constituency.

[23] The connection between the two can also be seen in the German rhotic, which varies in realisation between coronal and velar/uvular.

coronals). Thirdly, there is nothing in particular that labials and coronals share to the exclusion of velars.[24]

In this system, coronals have many empty positions, and (part of) all that empty structure is available to adjacent segments: Words like *fiend* (long vowel plus cluster) are possible because the vowel "borrows" space from coronals (Pöchtrager 2010). The big size of coronals also explains why, say, tapping in English targets coronal plosives, but not labials or velars: They are the biggest objects in the system, and lenition is simply the removal of structure, which starts with the bigger segments before affecting the smaller ones (Pöchtrager 2016).[25] In addition to that, there are other aspects of those structures that can be exploited and that are directly relevant to the unification attempts of this article. I will turn to these in the next subsection.

3.2. Lowering before rhotics

The European French rhotic **ʁ** requires that preceding tautosyllabic mid vowels are open-mid (**ɛ œ ɔ**), not close-mid (**e ø o**). (As the wording suggests, this is first and foremost about distribution and does not necessarily imply any kind of derivation from close-mid to open-mid.) Crucially, **ʁ** has no such effect on high vowels (Tranel 1987): *mer* **mɛːʁ**/***meːʁ** 'sea', *heure* **œːʁ**/***øːʁ** 'hour', *or* **ɔːʁ**/***oːʁ**, 'gold'.[26] This division amongst the vowels (high unaffected by lowering, mid affected) is of course reminiscent of SY and suggests that a similar solution should be sought. This in turn allows us to make headway into the representation of consonants.

I submit that **ʁ** contains a position Specxo which must also appear on a preceding vowel (again, not in a derivational sense), in much the same way as the Specxn of SY vowels was copied to the left. In this way, both the exclusion of close-mid vowels and the fact that high vowels are unaffected can be explained, as long as we make two (connected) assumptions: Firstly, French high and mid vowels have an internal structure like those of SY. Secondly, the requirement of a specifier to a higher head (in this case, xo) is only possible if the target also contains such a higher head, which in the case of vowels is xn. In other words, the specifier requires (once again) an appropriate foundation (cf. section 2.4). Mid vowels provide that foundation, high vowels do not. This suggests that higher heads (xo, xn) and lower heads (xO, xN) each form a natural class whose members stand in a special relationship to one another, such that the specifier of one is compatible with the other. Note in addition that French does not limit elements to the lower head xN, unlike SY in (12): French does have front rounded vowels. The argument for French

[24]Velars and labials being united by small size makes the assumption that both classes contain **U** (as e.g. in Backley 2011) unnecessary.
One reviewer worries that the structural reinterpretation of elements means that the acoustic signature associated with elements (Harris & Lindsey 1995) is therefore given up. This worry is unfounded; the theory does subscribe to identifiable acoustic signatures but it simply claims that they are associated with structures as well as elements.

[25]Plosives are also bigger than fricatives or approximants in that system (Pöchtrager 2006), thus coronal plosives end up as the biggest.

[26]The rhotic **ʁ** is not the only lowering agent in European French, there are other contexts that are conducive to lowering, often having to do with constituent structure (the "loi de position"). But tautosyllabic **ʁ** is the most reliable lowering agent in French: "The closing of a syllable by the consonant [r] thus plays an absolute role in the distribution of the mid vowels: it forces in all cases the appearance of the half-open mid vowels, destroying the potential influence of all the other factors which might tend to make the half-closed mid vowels occur" (Tranel 1987, 59). Surely any analysis of vowel distribution should start with the most reliable contexts.

can be directly applied to yet another language, Swedish, where lowering before the rhotic ɹ is also observed, yet again affecting mid vowels but leaving high vowels unaffected (Engstrand 1999, 141).[27]

The lowering data suggest that those rhotics contain a Specxo. In French and Swedish at least, lowering is specific to rhotics and does not extend to other consonants in the system, suggesting that only rhotics contain a Specxo.[28] This difference raises the question why that should be so. Why would rhotics need a Specxo while other consonants can do without? This issue will also come up again in the next subsection. So far the answer is only partially clear, but let me give a general discussion.

The Swedish rhotic is coronal. Recall from the previous section that in GP 2.0 both coronals (and velars) are pure structure. (They only differ in the number of heads.) It is tempting to blame the lowering behaviour on the (generally) bigger size of coronals. Diachronic change shows that also (the decomposition of) *s*, another coronal, can lead to lowering: In many varieties of Spanish *s* has debuccalised to *h*, which in turn has led to the lowering of an adjacent vowel (Hernández-Campoy & Trudgill 2002). If both coronals and velars (where I subsume *h*) are pure structure and differ only in the number of heads, then debuccalisation of *s* to *h* simply means that *s* has lost the relevant structure making it coronal, and (part of) that structure has moved over to the vowel, leading to lowering.[29] That *s* should show behaviour similar to rhotics is somewhat unsurprising, at least given their affinity in diachronic change (rhotacism). However, as pointed out before, the lowering behaviour cannot even be generalised to all coronals in Swedish. And of course French with its dorsal[30] rhotic precludes generalisation to coronals even more clearly.

There is one other thing that both the Swedish and the French rhotic share (besides being a rhotic, that is): They are both devoid of melody, since, once again, this is what unites coronals and velars (in the broadest sense). While it is unclear (and indeed rather unexpected) that being without melody implies having a Specxo, section 3.4 on emphatics will support the conclusion that there is a connection.

3.3. Retroflexes

Given that the theory does allow consonants with Specxo (as postulated for rhotics), the question arises what it would mean for other consonants besides rhotics to have a Specxo. I submit that Specxo also encodes retroflexion (to be discussed in this subsection) and emphaticness (next subsection). Consider the data in (22), taken from Riad (2013, 73ff, transcription of vowel length modified to reflect surface realisation), which show how Swedish ɹ and a following alveolar merge as a (single) retroflex.

[27]It is unclear to me to what extent Québec French high vowel lowering (Walker 1984; Pöchtrager 2018b) can be subsumed under this. That process singles out high vowels (**vide** 'to empty' vs. **vɪd** 'empty'), but is subject to slightly different triggering conditions. Some varieties allow lowering to be passed on to vowels to the left (Poliquin 2006), reminiscent of harmony in SY.

[28]Or, if other consonants do contain Specxo, they do not pass it on.

[29]I am grateful to Maria-Rosa Lloret (p.c.) for making me aware of this connection.

[30]The characterisations of French *r* vary widely: dorsal, velar, uvular, pharyngeal (Tranel 1987, 142).

(22) a. **stuːɹ** + **t** → **stuːʈ** 'big N.'
b. **fœːɹ** + **d** → **fœːɖ** 'led PAST PART.'
c. **ˈlɑːgɛɹ** + **n** → **ˈlɑːgɛɳ** 'the laurel'

The lowered mid-vowel in (22b) shows that the quality of the vowels preceding **ɹ** is as expected. But why does the following coronal change into a retroflex?[31] There is (the sketch of) a GP analysis in Backley (2011, 95), who assumes that all of **t d s n l** *and* **ɹ** contain **A** as a non-head (which for him means they are alveolar), but when **ɹ** merges with one of the other consonants, **A** is promoted to head, which makes the consonant retroflex. Since all of those consonants start out with **A** as non-head, this leaves unaddressed at least two issues: (i) Why should something happen only when **ɹ** comes to stand next to one of the other consonants, instead of any other two consonants out of the total set? (ii) Why should **A** be promoted to head in that particular context?

The alternative I would like to suggest answers both questions at the same time and capitalises on the lowering properties of **ɹ** discussed in the previous section: Retroflexes are simply coronals with an additional Specxo. More concretely, while Swedish **ɹ** contains Specxo, the other consonants involved in the merger do not. When **ɹ** merges with one of them, the only thing that remains of **ɹ** is its Specxo, which survives in the output consonant. Crucially, the presence of Specxo not only explains lowering of preceding (mid) vowels, but at the same time serves to express retroflexion. This is of course expected if lowering and retroflexion are taken to be the same phonologically. (Retroflexes could be called "lowered" coronals.) Since Specxo remains also after the merger, mid vowels preceding **ɹ** or one of the retroflexes will have (and keep) their typical lowered quality. As for the merger itself, no arbitrary promotion of an element's status is required, and the special status of **ɹ** falls out.[32]

3.4. Emphatics

Many Semitic languages display emphatic consonants (phonetically: pharyngealised, velarised), which "impose a backing effect on surrounding consonants and backing and lowering effects on vowels in their neighborhood" (Mustafawi 2018, 24). In terms of (binary) distinctive features, this is usually analysed as [+RTR]-spread (Watson 1999; Davis 1995), comparable to [−ATR].[33] In the framework advocated here this can again be analysed as the copying of Specxo, that the lowering of the vowel is a visible sign of. What this implies is that retroflexes and emphatics are essentially the same. This unification also seems plausible when looking at the place of articulation of emphatics: Emphatics are typically coronal; non-coronal emphatics are often derived ("allophonic") or altogether debated, cf. Mustafawi (2018, 15f). This is unsurprising if coronals are taken to contain two heads (cf. section 3.1), xo and xO, as that will guarantee that the higher head xo is present and can serve as the appropriate foundation for Specxo.

[31] The phonetic characterisation of Swedish consonants varies somewhat. Engstrand (1999, 140) classifies **ɹ** as alveolar and **t d s n l** as dental. Riad (2013, 45, 68), on the other hand, refers to **ɹ** as phonetically alveolar but phonologically retroflex, while **t d s n l** are given as dental *or* alveolar.

[32] Important questions remain for future research, e.g. what happens to the position given up by **ɹ**?

[33] Comparable or identical depending on whether [+RTR] is seen as only similar to [−ATR] or in fact identical.

In fact, this bold equation of emphatics and coronals requires a slight toning-down, and this takes us back to an issue that came up when looking at the lowering before rhotics in section 3.2. For one thing, Arabic **q** is the emphatic counterpart to **k g**, just as **ṭ** is to **t d** (McCarthy 1994; Watson 2002), suggesting that emphaticness is not inextricably bound to coronality. Similarly, in Quechua, **q** $\mathbf{q^h}$ **q' h** lower adjacent high vowels to mid (Cusihuamán 2001, 6) (though apparently not completely exceptionlessly). If Specxo is the cause of lowering, it must be included in those consonants, too; and therefore there has to be a higher head xo, but crucially without the lower head xO. (The presence of both heads would yield a coronal.) Emphaticness is not exclusive to coronals, as long as the correct head is present (just as [−ATR] is not absolutely exclusive to open vowels, but more likely). For some reason, however, specifiers to the higher head seem to favour structures devoid of melody, i.e. coronals and velars in the case of consonants, and more open vowels in the case of vowels. The reason for this particular connection remains unclear for now.

4. JAKOBSON AND SCHANE REVISITED

As mentioned before, a wholesale merger of categories as far as their phonological representation is concerned has been attempted before: Schane (1990) argues that aperture, [−ATR], and laxness should all be subsumed under the (privative) particle **a**, while Jakobson's feature [flat] covers rounding, retroflexion, velarisation, and pharyngealisation (Jakobson, Fant & Halle 1969). While I am obviously standing on the shoulders of giants, since the account presented here owes a tremendous debt to those sources, there are still some crucial differences to and original contributions by the present article that I would like to throw into relief in this section.

First and foremost, Schane's account in the framework of Particle Phonology dealt with vowels only. It remains unclear how and to what extent his account can be extended to consonants. There has been work in Particle Phonology arguing that **a** also covers coronality (Broadbent 1996, 1999), in the same way as has been done for the element **A** in GP (cf. section 2.3). But coronality is not the central focus of the present article (though I do of course subscribe to the claim that **A**, or rather its successor in the form of empty structure, should cover coronality). What is argued here is that *other* consonantal properties, like retroflexion and emphaticness, need to be thrown into the mix, and combined with vocalic qualities like aperture and [ATR].

My account of the diachronic development of high [−ATR] vowels (in section 2.5) is in one aspect identical to Schane's: Like him, I assume that a change from [−ATR] to additional openness can take place so easily because those two properties are really the same thing: For him the particle **a**, for me empty structure. In Schane's case the role played by **a** (aperture or [−ATR]) is determined by whether the particle resides in the "core" of a segment or on a separate ATR-tier. But the role (and number) of tiers in Schane's system is never really discussed in detail; a tier dedicated to a particular property (ATR in this case) only comes up at this point and nowhere else. In contrast, in my account the internal organisation of a nucleus is explicitly given and has been put to the test in other languages and for other phenomena, cf. the references in section 2.3. Also, in Schane's account it remains unclear why such an exchange of particles should happen—it just does. In contrast, in my system with two nuclear heads some motivation can be be given for why a change should happen. (As always in diachrony, there is of course no way of predicting *that* something will happen.)

There is one aspect where Schane's account is superior to the one presented here, at least at this point in the development of GP 2.0: His particle **a** covers not only aperture and [−ATR], but also laxness. This allows him to account for why lax high vowels lowered by one degree to close-mid on the way from Old to Middle English (Schane 1990, 5f), with an entirely similar case in Vulgar Latin (Schane 1984, 145): Since **a** expresses both laxness and openness, the exchange can be seen as entirely parallel to what happened to high [−ATR] vowels. Within GP 2.0, laxness sometimes does involve empty structure, but not consistently so (Pöchtrager 2020), which is why it is unclear to what extent Schane's analysis could be translated without problems.[34]

In contrast to Schane, Jakobson explicitly designed his features to be generally applicable to both consonants and vowels. Jakobson's feature [flat] covers rounding, retroflexion, velarisation, and pharyngealisation. That particular unification receives at least some partial support from phoneticians since "[t]here is some similarity in quality between retroflex stops and velarized or pharyngealized stops, because in all these sounds, the front of the tongue is somewhat hollowed" (Ladefoged & Johnson 2010, 236). The feature [flat] illustrates of course a general strategy in Jakobson's work, which "assimilates as many traditional phonetic dimensions as possible to one another, […] whenever […] they cannot be shown to function independently of one another" (Anderson 1985, 122). But there is some important criticism, since while the merger might generally be reasonable, it has been argued that certain necessary distinctions are lost. For example, Ubykh and Abkhaz have been brought up as counterexamples where independent contrasts of retroflexion and rounding are reported (Anderson 1985, 124). This raises the question whether my account repeats old mistakes and whether I am biting off more than I can chew?

In fact, it is probably no accident that it is the independence of rounding in Ubykh and Abkhaz that has been brought up. Rounding also figures prominently in a criticism against Jakobson's [flat] raised by McCawley (1967). The argument runs roughly as follows: Arabic emphatics are [flat]. Arabic **u** is round, therefore also [flat]. Vowels following emphatics are pharyngealised, i.e. the feature [flat] is copied over. I will simply call this "[flat] spreading" for ease of reference. What is the proper (phonetic) interpretation of [flat] then? In consonants, [flat] must mean pharyngealised, while in vowels that are high and back it must (in general) mean round. A problem arises when a pharyngealised consonant is followed by a high back vowel: Such a vowel will be [flat] in any case, but how do we know that [flat] in this case means both rounding *and* pharyngealisation, instead of *only* rounding? The only way to guarantee this is to assume that vowels adjacent to pharyngealised consonants are themselves realised as pharyngealised, but then we duplicate the statement of [flat] spreading, by which [flat] is redundantly assigned to vowels adjacent to [flat] consonants.

The question is now, have I fallen into the same trap and lost distinctions that need to be maintained? The properties covered by Specxo (and Specxn) are similar to what is captured by [flat]. Crucially, rounding is *not* amongst those properties. Rounding is separate and, like in other GP approaches, mediated by the element **U**. Most of Jakobson's problems with [flat] stem from the inclusion of rounding. But lumping together retroflexion, velarisation, and

[34] Whether an **a** is interpreted as laxness or aperture (with accompanying differences in behaviour, e.g. distribution in open/closed syllable) cannot be read off Schane's representations directly, but must be assessed from the system as a whole, i.e. what other vowels a given vowel contrasts with. This is quite unlike the case of the ATR-tier, which allows an unequivocal identification.

pharyngealisation (as done by Specxo) seems less of a problem, just like the merger of aperture and [−ATR]) (as done by Specxn) or the unification of both (as a specifier to a higher head).

Where this article goes well beyond what is covered by either Jakobson or Schane, is of course in the claim that both those strands of research should be combined. Furthermore, by positing that it is a structural property underlying the various phonetic realisations, not a melodic one (be it the particle **a** or the feature [flat]), I also capture various length-related phenomena that were the original motivation for replacing **A** by structure in the first place.

5. SOME OPEN QUESTIONS AND A CONCLUSION

This article has tried to show that appearances can be deceiving and that what we traditionally treat as different phonetic properties (aperture, [ATR], retroflexion, emphaticness) are phonologically very much the same, in that they all involve empty structure. This is particularly clear when looking at nuclei, where not only the phonetic realisation is comparable, but where seemingly "different" properties can stand in for one another: Empty skeletal positions in nuclei are an expression of openness, whether that be openness as understood when we talk about **i** vs. **e**, or the kind of openness brought about by [−ATR]. (All of them have the phonetic effect of raising F_1.) In that sense, [−ATR] is simply a specialised term for one *kind* of openness, and the only reason why we need a special term is because we want to distinguish between different subtypes of openness that are relevant for phonological classes. Diachronic changes where one type is exchanged for the other ([−ATR] being reinterpreted as in section 2.5) support the view that, despite the existence of subclasses, both [−ATR] and openness revolve around the same phonological objects: empty skeletal positions. As argued throughout the text, that kind of unification of seemingly disparate properties is not limited to nuclei.

As always, many open questions remain: For example, [−ATR] is structurally more complex than [+ATR], in that the former involves a specifier, unlike the latter. This might seem surprising (and even contradictory) if one goes by a traditional assumption that [−ATR] is the less marked member of the two, cf. e.g. Archangeli & Pulleyblank (1989, 204) and references given there. But there is only a contradiction if one assumes that structural size can be equated with markedness. Let me stress that this is *not* the position of GP 2.0. In fact, equating structural size with markedness never seems to hold. For example, stops are also bigger in size than fricatives in this theory (Pöchtrager 2006), yet stops are certainly less marked than fricatives, if anything. Similarly, an open vowel will have more structure than a high vowel, and again, open **a** is hardly marked. What the structures of GP 2.0 do seem to correlate with more clearly is what Harris (2006) refers to as the extent of deviation from the carrier signal; they contribute to the amount of information that is contained in the speech signal.

The question about markedness becomes more pressing when looking at so-called dominant/recessive systems of harmony (Halle & Vergnaud 1981; Lodge 1995), typically analysed as involving the spreading of [+ATR]. If [−ATR] can be passed on by copying a specifier, then how could [+ATR] be passed on? One potential way out is to follow Harris & Lindsey (1995) or Denwood (2002), and to assume that properties in recessive position can also be lost unless licensed by a corresponding property in dominant position. In such a scenario, [−ATR] in recessive position would be lexically present, but only survive to the surface if there was another [−ATR] in the dominant nucleus that could license it. This would "emulate" spreading of [+ATR], but future research will have to show if this is the most satisfactory way to go about it.

REFERENCES

Anderson, Stephen R. 1985. Phonology in the twentieth century: Theories of rules and theories of representations. Chicago, IL: University of Chicago Press.

Archangeli, Diana and Douglas Pulleyblank. 1989. Yoruba vowel harmony. Linguistic Inquiry 20(2). 173–217.

Backley, Phillip. 1997. Monovalency and the status of RTR. UCL Working Papers in Linguistics 9. 337–360.

Backley, Phillip. 2011. An introduction to Element Theory. Edinburgh: Edinburgh University Press.

Bamgbọṣe, Ayọ. 1967. Vowel harmony in Yoruba. Journal of African Languages 6. 268–273.

Basbøll, Hans and Johannes Wagner. 1985. Kontrastive Phonologie des Deutschen und Dänischen: Segmentale Wortphonologie und -phonetik. Tübingen: Max Niemeyer Verlag.

Broadbent, Judith M. 1991. Linking and intrusive r in English. UCL Working Papers in Linguistics 3. 281–301.

Broadbent, Judith M. 1996. The representation of coronal segments. Doctoral dissertation. University of London, London.

Broadbent, Judith M. 1999. A new approach to the representation of coronal segments. In S. J. Hannahs and M. Davenport (eds.) Issues in phonological structure: Papers from an international workshop. Amsterdam & Philadelphia, PA: John Benjamins. 1–25.

Charette, Monik and Aslı Göksel. 1996. Switching and vowel harmony in Turkic languages. In H. Kardela and B. Szymanek (eds.) A Festschrift for Edmund Gussmann from his friends and colleagues. Lublin: The Catholic University of Lublin Press. 29–55.

Clements, George N. 1985. Akan vowel harmony: A nonlinear analysis. In D. L. Goyvaerts (ed.) African linguistics: Essays in memory of M. W. K. Semikenke. Amsterdam & Philadelphia, PA: Benjamins. 55–98.

Cobb, Margaret. 1997. Conditions on nuclear expressions in phonology. Doctoral dissertation. School of Oriental and African Studies, University of London, London.

Cobb, Margaret. 2003. Government phonology and the vowel harmonies of natal Portuguese and Yoruba. In S. Ploch (ed.) Living on the edge: 28 Papers in honour of Jonathan Kaye. Berlin & New York, NY: Mouton de Gruyter. 223–242.

Crothers, John. 1978. Typology and universals of vowel systems. In J. H. Greenberg, C. A. Ferguson and E. A. Moravcsik (eds.) Universals of human language, Vol. 2: Phonology. Stanford, CA: Stanford University Press. 93–152.

Cusihuamán, Antonio G. 2001. Gramática Quechua Cuzco-Collao. Cuzco: CBC.

Cyran, Eugeniusz. 1997. Resonance elements in phonology: A study in Munster Irish. Lublin: Wydawnictwo Folium.

Davis, Stuart. 1995. Emphasis spread in Arabic and grounded phonology. Linguistic Inquiry 26(3). 465–498.

den Dikken, Marcel and Harry van der Hulst. 2020. On some deep structural analogies between syntax and phonology. In K. Nasukawa (ed.) Recursion in phonology. Berlin & New York, NY: Mouton de Gruyter. 57–115.

Denwood, Ann. 2002. Vowel harmony in government phonology: Three for the price of one. Journal of the Altaic Society of Korea 12. 155–182.

Dresher, B. Elan. 2009. The contrastive hierarchy in phonology. Cambridge: Cambridge University Press.

Engstrand, Olle. 1999. Swedish. In International Phonetic Association (ed.) Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet. Cambridge: Cambridge University Press. 140–142.

Fudge, Erik C. 1969. Syllables. Journal of Linguistics 5. 253–286.

Goh, Yeng-Seng. 1997. The segmental phonology of Beijing Mandarin. Taipei: Crane Publishing Co., Ltd.

Halle, Morris and Jean-Roger Vergnaud. 1981. Harmony processes. In W. Klein and W. Levelt (eds.) Crossing the boundaries in linguistics: Studies presented to Manfred Bierwisch. Dordrecht, Boston, MA & London: Reidel. 1–22.

Harris, James W. 1983. Syllable structure and stress in Spanish. Cambridge, MA & London: MIT Press.

Harris, John. 1990. Segmental complexity and phonological government. Phonology 7(2). 255–301.

Harris, John. 2006. The phonology of being understood: Further arguments against sonority. Lingua 116. 1483–1494.

Harris, John and Geoff Lindsey. 1995. The elements of phonological representation. In J. Durand and F. Katamba (eds.) Frontiers of phonology: Atoms, structures, derivations. London & New York, NY: Longman. 34–79.

Hernández-Campoy, Juan Manuel and Peter Trudgill. 2002. Functional compensation and Southern Peninsular Spanish /s/ loss. Folia Linguistica Historica 23(1–2). 31–57.

Jakobson, Roman, C. Gunnar Fant and Morris Halle. 1969. Preliminaries to speech analysis: The distinctive features and their correlates, 9th printing edn. Cambridge, MA: The MIT Press.

Kaye, Jonathan. 1980. The mystery of the tenth vowel. Journal of Linguistic Research 1. 1–14.

Kaye, Jonathan, Jean Lowenstamm and Jean-Roger Vergnaud. 1990. Constituent structure and government in phonology. Phonology 7(2). 193–231.

Kaye, Jonathan and Markus A Pöchtrager. 2013. GP 2.0. SOAS Working Papers in Linguistics & Phonetics 16. 51–64.

Ladefoged, Peter and Keith Johnson. 2010. A course in phonetics, 6th edn. Boston, MA: Wadsworth.

Lindau, Mona. 1978. Vowel features. Language 54(3). 541–563.

Lodge, Ken. 1995. Kalenjin phonology and morphology: A further exemplification of underspecification and non-destructive phonology. Lingua 96. 29–43.

McCarthy, John J. 1994. The phonetics and phonology of Semitic pharyngeals. In P. Keating (ed.) Phonological structure and phonetic form. Cambridge: Cambridge University Press. 191–233.

McCawley, James D. 1967. Le rôle d'un système de traits phonologiques dans une théorie du langage. Langages 8. 112–123.

Mustafawi, Eiman. 2018. Arabic phonology. In E. Benmamoun and R. Bassiouney (eds.) The Routledge handbook of Arabic linguistics. London & New York, NY: Routledge. 11–50.

Ọla, Ọlanike. 1992. Yoruba vowel harmony. MA thesis. SOAS, University of London, London.

Paradis, Carole and Jean-François Prunet (eds.). 1991. Phonetics and phonology, Vol. 2: The special status of coronals: Internal and external evidence. San Diego, CA: Academic Press.

Pöchtrager, Markus A. 2006. The structure of length. Doctoral dissertation. University of Vienna, Vienna.

Pöchtrager, Markus A. 2010. The structure of A. Paper presented at the 33rd GLOW Colloquium, 13–16 April 2010, Wrocław.

Pöchtrager, Markus A. 2012. Deconstructing A. Paper presented at the MFM Fringe Meeting on Segmental Architecture, 23 May 2012, University of Manchester, Manchester.

Pöchtrager, Markus A. 2013. On A. Paper presented at the Workshop on Melodic Representation, 12 March 2013, UCL, London.

Pöchtrager, Markus A. 2016. It's all about size. In P. Szigetvári (ed.) 70 snippets to mark Ádám Nádasdy's 70th birthday. http://seas3.elte.hu/nadasdy70/pochtrager.html.

Pöchtrager, Markus A. 2017. Transparent vowels: Small cogs in large machines. Paper presented at the 25th Manchester Phonology Meeting, 25–27 May 2017, University of Manchester, Manchester.

Pöchtrager, Markus A. 2018a. Sawing off the branch you are sitting on. Acta Linguistica Academica 65(1). 47–68.

Pöchtrager, Markus A. 2018b. Tense? (Re)lax! A new formalisation of the tense/lax contrast. Paper presented at the 26th Manchester Phonology Meeting, 24–26 May 2018, University of Manchester, Manchester.

Pöchtrager, Markus A. 2020. Tense? (Re)lax! A new formalisation for a controversial contrast. Acta Linguistica Academica 67(1). 53–71.

Pöchtrager, Markus A. 2021a. English vowel structure and stress in GP 2.0. In S. Bendjaballah, A. Tifrit and L. Voeltzel (eds.) Perspectives on element theory. Berlin: Mouton de Gruyter. 157–183.

Pöchtrager, Markus A. 2021b. Towards a non-arbitrary account of affricates and affrication. Glossa: A Journal of General Linguistics 6(1). 1–31. https://doi.org/10.5334/gjgl.1116.

Pöchtrager, Markus A. 2023. The unbearable lightness of being high: Openness as structure and the consequences for prosody. In J. van de Weijer (ed.) Representing phonological detail, Part I: Segmental structure and representations. Berlin: De Gruyter. 71–89.

Pöchtrager, Markus A. 2024. Vowel reduction in European Portuguese and the removal of structure. Isogloss 10(3)/2. 1–22.

Pöchtrager, Markus A. Submitted. Transparent vowels: Small cogs in large machines. Acta Linguistica Academica.

Polgárdi, Krisztina. 1998. Vowel harmony. An account in terms of government and optimality. The Hague: Holland Academic Graphics.

Poliquin, Gabriel Christophe. 2006. Canadian French vowel harmony. Doctoral dissertation. Harvard University, Cambridge, MA.

Rennison, John. 1990. On the elements of phonological representations: The evidence from vowel systems and vowel processes. Folia Linguistica 24(3–4). 175–244.

Riad, Tomas. 2013. The phonology of Swedish. Oxford: Oxford University Press.

Schane, Sanford A. 1984. The fundamentals of particle phonology. Phonology Yearbook 1. 129–155.

Schane, Sanford A. 1990. Lowered height, laxness, and retracted tongue root: Different manifestations of phonological aperture. Word 41(1). 1–16.

Selkirk, Elisabeth O. 1982. The syllable. In H. van der Hulst and N. Smith (eds.) The structure of phonological representations, Vol. II. Dordrecht & Cinnaminson, NJ: Foris Publications. 337–383.

Stewart, John M. 1971. Niger-Congo, Kwa. In T. A. Sebeok (ed.) Current trends in linguistics 7: Linguistics in Sub-Saharan Africa. The Hague: De Gruyter Mouton. 179–212.

Tranel, Bernard. 1987. The sounds of French. Cambridge: Cambridge University Press.

Vaux, Bert and Andrew Wolfe. 2009. The appendix. In E. Raimy and C. E. Cairns (eds.) Contemporary views on architecture and representations in phonology. Cambridge, MA: MIT Press. 101–143.

Walker, Douglas C. 1984. The pronunciation of Canadian French. Ottawa: University of Ottawa Press.

Watson, Janet C. E. 1999. The directionality of emphasis spread in Arabic. Linguistic Inquiry 30(2). 289–300.

Watson, Janet C. E. 2002. The phonology and morphology of Arabic. Oxford: Oxford University Press.

Živanović, Sašo and Markus A. Pöchtrager. 2010. GP 2.0 and Putonghua, too. Acta Linguistica Hungarica 57(4). 357–380.