3 x Phonology

TOBIAS SCHEER

Université Côte d'Azur, CNRS, Nice, France
scheer@unice.fr

Abstract

What does it take to run a full substance-free phonology (SFP)? Because in classical approaches only items below the skeleton have phonetic properties that, according to SFP, need to be expunged, current work in SFP only ever concerns segmental aspects. If substance is removed from segmental representation and primes and computation are therefore arbitrary, the non-trivial question arises: how can such a system communicate with a system where primes and computation are not arbitrary (at and above the skeleton)? The two phonologies below and at / above the skeleton that exist in production are complemented with a third phonology that occurs upon lexicalization, that is, when L1 learners or adults transform the acoustic signal into a stored representation. The core of this article argues that this broad architecture is inhabited by three distinct computational systems along the classical feature geometric divisions: Son(ority) is located at and above the skeleton, while Place and Lar(yngeal) live below the skeleton. The question then is how a multiple-module spell-out works, that is, how ingredients from three distinct vocabularies can be mapped onto a single phonetic item. It is argued that the skeleton plays a central role in this conversion.

Keywords: modularity, spell-out, sonority, lexicalization, substance-free

Résumé

Quels ingrédients faut-il pour qu'une phonologie sans substance (PSS) puisse fonctionner ? Dans les approches classiques, seuls les éléments situés sous le squelette ont des propriétés phonétiques qui, selon la PSS, doivent être éliminées. C'est la raison pour laquelle les travaux à ce jour en PSS ne concernent que ces aspects segmentaux. Si la substance est supprimée de la représentation segmentale et que les primitives et la computation sont donc arbitraires, la question non triviale se pose de savoir comment un tel système peut communiquer avec un système où les primitives et la computation ne sont pas arbitraires (squelette et au-dessus). Les deux phonologies (en-dessous et au-dessus du squelette) qui existent en production sont complétées par une troisième phonologie qui opère lors de la lexicalisation, c'est-à-dire lorsque les apprenants L1 ou les adultes transforment le signal acoustique en une représentation stockée. Le noyau de cet article soutient que cette architecture globale est habitée par trois systèmes computationnels distincts selon les divisions classiques de la géométrie de traits : Son(orité) est situé au niveau et au-dessus du squelette, tandis que Place et Lar(yngal) se situent en dessous du squelette. La question est alors de savoir comment fonctionne un *spell-out* (épel) à modules multiples, c'est-à-dire, comment les

ingrédients de trois vocabulaires distincts peuvent correspondre à un seul item phonétique. Il est avancé que le squelette joue un rôle central dans cette conversion.

Mot clés: modularité, *spell-out* (épel), sonorité, lexicalisation, sans-substance

1. INTRODUCTION[1]

The goal of this article is to make explicit what it takes to run a full substance-free phonology (SFP).[2] By its very essence, SFP has concentrated on the locus of substance in phonology, (i.e., the area below the skeleton). The area at and above the skeleton is not studied, or is understudied in this framework, and as far as I can see the relationship between these areas has not received any attention thus far. It is shown below that substance-free melodic primes raise non-trivial questions for how the two areas communicate: familiar syllabification algorithms, for example, cannot work with substance-free primes.

The idea of phonology being substance-free entails i) that any substance which is present in current approaches needs to be removed and ii) that there is a place where substance can and should exist, which is not the phonology. Substance being another word for phonetic properties, ii) means that SFP is necessarily modular in kind: there is a phonological system that is different from the phonetic system. They communicate but do not overlap: there is no phonology in phonetics and no phonetics in phonology. This describes exactly the Fodorian idea of modularity (on which more in section 2.4.2) where distinct computational systems work on domain-specific sets of vocabulary and communicate through a translational device (spell-out).

The consequence of i) is the starting point of the article: to date work on SFP only ever concerns melody, that is, items which in a regular autosegmental representation occur below the skeleton. The reason is that this is where phonetic properties are located and need to be removed: there is nothing to be removed from items at and above the skeleton because they have no phonetic properties. That is, the feature [±labial] has phonetic content, but an onset, a grid mark or a prosodic word do

---

[1]The paper is a result, among other things, of the discussion with local phonologists in Nice: Diana Passino, Alex Chabot and Paolo Danesi.

[2]A reviewer points out that it is also useful to make explicit what *substance* means in the substance-free literature, to avoid misunderstandings. I take his definition: "any non-symbolic, non-phonological aspect of realization – either any reference to human anatomy, to acoustic properties, or behaviour (perceptibility, salience, confusability etc.)."

It may also be worth warning the reader that this article does nothing to motivate SFP, to explain its *raison d'être*, or to compare its merits with respect to other approaches. Rather, SFP is taken for granted and its consequences for the workings of phonology outside of the melodic realm are explored. *Abbreviations used*: CSE: Cochlea-scaled Spectral Entropy; DP: Derivational Phonology; ERB: Equivalent Rectangular Bandwidth; DP: Derivational Phonology; GP: Government Phonology; Lar: Laryngeal; LF: logical form; LSP: language-specific phonetic system; MSC: Morpheme Structure Constraints; OT: Optimality Theory; PF: phonetic form; SFP: substance-free phonology; Son: Sonority; SPE: *Sound Patters of English*.

not. This is why work in SFP never talks about syllable structure or other items that occur at and above the skeleton (more on this in section 2.1).

Therefore, what it takes to make SFP a viable perspective is not only to remove substance from phonology, to show how phonology relates to phonetics, and to explain how acquisition works: it is also necessary to answer a number of non-trivial questions that arise when the area from which substance was removed and where primes and computation are therefore arbitrary (below the skeleton) communicates with a system where primes and computation are not arbitrary (at and above the skeleton). This distinction is the heart of the article (section 2).

The two phonologies at hand are active in speech production, that is, when a multi-morphemic string (a phase, or cycle) that was pieced together from items stored in long-term memory (and which also features phonological representatives of morpho-syntactic divisions, such as # in SPE) is submitted to phonological interpretation. Phonological activity also occurs prior to production, though: in order for a lexical item to be stored in long-term memory, the gradient and non-cognitive acoustic signal needs to be converted into a discrete cognitive representation. Lexicalization phonology (as opposed to production phonology) is introduced in section 3. This then amounts to three phonologies altogether (two in production, one upon lexicalization), which is one motivation for the title of the article.

Sections 4 and 5 take a closer look at the content of the two areas that are distinguished in production. It is argued in section 4 that the traditional division of the area below the skeleton into sonority, place and laryngeal primes[3] (as embodied in Feature Geometry) in fact defines three distinct computational systems (modules): Son, Place and Lar.[4] In section 5, these are shown to either live at and above (Son) or below the skeleton (Place, Lar). A lexical entry of a segment has thus three compartments hosting items of three distinct vocabulary sets (just as the lexical entry of a morpheme is made of three distinct vocabulary sets: morpho-syntax, phonology, semantics). The challenge then is to explain how the three modules communicate in general, and what guarantees *segment integrity* in particular (how does the system "know" which Son primes, Place primes and Lar primes belong to the same segment?). In short, the question is how does a *multiple-module spell-out* work, that is, how ingredients from three distinct vocabularies can be mapped onto a single phonetic item.[5] It is argued that the skeleton plays a central role in this

---

[3]In this article, the generic term *prime* is used to refer to all (theory-)specific incarnations of items that occur below the skeleton, such as binary or monovalent features, Elements, particles etc. Differences between these items are irrelevant for the argument.

[4]Note that in this perspective there is nothing like a "sub-module", that is, a module in a module. There are only computational systems that communicate. That is, there is no phonology module that includes three sub-modules. If *phonology* is correctly characterized as a set of three distinct modules, there is no phonology module. *Phonology* is then just a handy word to refer to the set of the three computational systems mentioned (in the way chemists talk about *water* when they mean H_2O).

[5]Note that this is the mirror image of the situation at the upper interface of phonology with morpho-syntax where (in production) the vocabulary of a single module (morpho-syntax) is converted into the items of two distinct systems (phonology / PF, semantics / LF).

conversion. The division into three content-defined modules Son, Place, Lar also motivates the title of the article.

Section 5 also evaluates the consequences of the idea that Son primes are phonologically meaningful (non-arbitrary), while Place and Lar primes are phonologically meaningless (arbitrary). It is argued that the former match the Concordia version of SFP (Hale, Reiss, Kissock, Volenec, see section 5.2.4) where primes and their association to phonetic categories are given at birth, while the latter instantiate the take of all other SFP approaches where both primes and their association to phonetic categories are emergent (i.e., acquired by the infant).

Section 6 considers a pervasive question raised by the presence of a spell-out in addition to regular phonological computation: how can we know whether a given alternation is due to the former (interpretational), or to the latter (phonological)? Finally, the conclusion in section 7 addresses more general issues that arise given the distinctions discussed: the so-called duplication problem (does phonological computation do the same labour twice: upon lexicalization and upon production?), the wiring of modules (why does Son, but not Place or Lar, communicate with morpho-syntax?) and the specific status of Son primes as non-arbitrary and hard-wired properties present at birth (why do Place or Lar not have this privilege?).

2. WHAT EXACTLY IS SUBSTANCE-FREE?

This section introduces the basic distinction between phonological items below and above the skeleton: the former have a phonetic correlate, the latter do not; the former have arbitrary labels and may enjoy crazy computation (crazy rules), while the latter are drawn from a small set of cross-linguistically stable items (onset, nucleus etc.) which do not produce any crazy rules and whose labels are not interchangeable.

2.1. Phonological objects with and without phonetic correlates

The idea of substance-free phonology (SFP) concerns the melodic (or segmental) side of phonology, that is, items that are found below the skeleton in a regular autosegmental representation. The area at and above the skeleton is not within the purview of this approach.[6] The reason is that only items can be substance-free which are (wrongly, according to SFP) taken to have a substance. Substance is another word for phonetic properties, but onsets, rhymes, stress, prosodic words, metrical grids, skeletal slots or whatever other item occurs at and above the skeleton do not have any: only items below the skeleton (the world called segmental or melodic) have a phonetic correlate that SFP argues needs to be removed.

---

[6]As far as I can see, this holds for all versions of SFP or approaches that are akin to SFP (see the overview in Scheer 2019a), including Hale and Reiss (2000, 2008), Volenec and Reiss (2018), Boersma and Hamann (2008), Hamann (2011, 2014), Mielke (2008), de Carvalho (2002b), Odden (2006, 2022), Blaho (2008), Samuels (2012), Scheer (2014b), Iosad (2017), Dresher (2014, 2018), Chabot (2019).

Of course, items at and above the skeleton bear on the phonetic realization of segments (e.g., a lateral is pronounced [w] in a coda, but [ɫ] in an onset) – but they do not have any phonetic properties themselves. The absence of phonetic properties is obvious for onsets, nuclei, prosodic words, the metrical grid etc.[7] But it also holds true for prominence: the supra-skeletal item (ictus) is lexically defined or distributed according to an algorithm, and then represented above the skeleton (as foot structure, metrical grids, extra-syllabic space etc.) with no reference to its phonetic value. The phonetic correlates of prominence are not present in the phonological representation (unlike labiality in [±labial] etc.): they are only introduced in regular phonetic interpretation thereof. That is, prominent segments bear three phonetic properties (in all languages, Ladefoged and Ferrari-Disner 2012: 24): loudness (measured in decibels), duration (measured in milliseconds) and pitch (measured in Hertz), the latter two being able to be phonologized (as length or tone, respectively, see Hyman 2006). Therefore, there is no substance to be removed from the phonological representation of prominence, and SFP is not concerned with it.[8]

Note that the absence of phonetic correlates for items at and above the skeleton is not just an analytical choice that phonologists happen to have made when autosegmental representations were developed. Rather, it is a necessary property shared by all approaches that endorse items such as syllable structure on top of melodic building blocks, whether this is implemented in an autosegmental environment or not, and no matter whether melodic representations are substance-free or substance-laden. In substance-laden systems, primes carry their phonetic value in their name, while there is no such specification for items at and above the skeleton. In SFP, items below the skeleton (now in substance-free guise: α, β, γ etc.) will be specified for a phonetic correlate post-phonologically at the interface with phonetics (there are a number of ways this is done in the SFP literature; see the overview in Scheer 2019a): just like at the upper interface (of morpho-syntax with phonology, for example PAST TENSE ↔ -ed in English), spell-out is based on a list that infants need to acquire.

_____

[7]In Standard Government Phonology, empty nuclei (when ungoverned) are said to have a specific pronunciation, ɨ (or ə, Kaye 1990) – but this is the pronunciation of an *empty* nucleus, not of a nucleus. That is, the absence of melody is a melodic item, and it is this melodic item, zero, which has a phonetic correlate. Empty onsets in strong position are also sometimes said to be pronounced as a glottal stop (in German, see Alber 2001, or French, see Ségéral and Scheer 2001a: 116 footnote 18, Pagliano 2003), but these workings have no cross-linguistic generality (unlike what is assumed in Standard GP for ungoverned empty nuclei).
A reviewer suggests that items above the skeleton such as syllabic constituents may have phonetic correlates in terms of duration and amplitude in the acoustic signal. Duration and amplitude are not phonological in kind, though (unless of course duration is phonologized, i.e., becomes distinctive length): they are a matter of phonetic implementation (see section 5.6.1 on Language-Specific Phonetics).
[8]Tone has a phonetic correlate (pitch), shares properties of regular segmental features and interacts with them (Hyman and Schuh 1974), but, since autosegmental times is usually represented above the skeleton because of its "mobile" properties (relative autonomy with respect to the tone-bearing unit). See Hyman (2011: 206ff) and Iosad (2012: 41ff) for discussion. Further diagnostics for the (dual) status of tone appear in section 4.1.3.

This list (spell-out instructions) defines which phonological objects are associated to which phonetic categories, for example, α ↔ labiality, β ↔ backness etc. (Scheer 2014b). Certain phonological objects are part of this list, while others are absent, and the division is the one mentioned above: items below the skeleton (alphas, betas etc.) do entertain a spell-out relationship with a phonetic category, but items at and above the skeleton do not and hence are absent from this list. As was mentioned, they may influence the phonetic properties of segments (through phonological computation: ł → w in coda position, more on that in section 4.4), but are not themselves associated to any phonetic correlate (there is nothing like coda ↔ palatality etc., see section 6).

As far as I can see, the distinction between items that do (below the skeleton) and do not (at and above the skeleton) have a phonetic correlate plays no role in the literature and, being trivial, is not made explicit (Clements 2009: 165 is an exception). Considered from the phonetic viewpoint, it is about the presence or absence of a phonetic correlate. Looked at from the phonological perspective, it distinguishes between *meaningful* and *meaningless* items: the phonological prime to which, say, phonetic labiality is associated when the system is built (upon acquisition), is interchangeable.[9] It may be anything and its reverse as long as it is distinct from other primes: an alpha is not any more appropriate than a beta or a gamma. By contrast, an onset is not interchangeable with a nucleus: there is no way to replace one with the other. This is because items at and above the skeleton have stable cross-linguistic properties: nuclei host vowels, not consonants (except for syllabic consonants on a certain analysis), rhymes form a unit with the onset preceding, not following them, etc. They also belong to a universal inventory (all languages have nuclei etc.). By contrast, items below the skeleton do not belong to a cross-linguistically stable inventory, and they have no cross-linguistically stable properties. This is why the only way to name them is by arbitrary alphas, betas and gammas.[10] Also consider the fact that despite substantial effort, phonologists have never been able to establish a putative universal set of melodic primes (feature inventory) – but the inventory of items at and above the skeleton is trivial and undisputed (within a given theory: moras, onset, nuclei etc.).

Items with stable cross-linguistic properties are thus phonologically meaningful, while those that are randomly interchangeable because they do not have any stable cross-linguistic properties are phonologically meaningless. Another way to put things is to say that because items at and above the skeleton belong to a universal inventory and enjoy universal properties that cannot be derived from the environment, they are innate. By contrast, items below the skeleton are not innate: the child is born with the (domain-general) ability to categorize and is predisposed to

---

[9]But, of course, once an association of a phonological prime with a phonetic correlate is installed in a given language and acquired by learners, the prime at hand becomes phonologically meaningful: alpha will bear an identity distinct from beta because their phonetic correlates are distinct.

[10]This view is shared by all versions of SFP (see footnote 6), except the one developed at Concordia / Montreal: see section 5.2.4.

use this faculty in order to create phonological primes, that is, to reduce the acoustic signal to discrete phonological units (see Odden 2022). That is, the child knows at birth that phonology exists and requires the construction of domain-specific primes to be put to use (more on this in sections 5.2.1 and 5.2.2).

Text under (1) sums up the discussion thus far, showing that SFP in fact operates a fundamental segregation between two types of phonology, below and (at and) above the skeleton.

(1)   Phonologically meaningful and meaningless items

	items below the skeleton	items at and above the skeleton
phonologically	meaningless	meaningful
interchangeable	yes	no
phonetic correlate	yes	no
cross-linguistically stable inventory	no	yes
cross-linguistically stable properties	no	yes
present at birth (innate)	no	yes

2.2.   Consequences of being phonologically meaningless or meaningful

Removing phonetic substance from (melodic) primes turns them into naked place-holders of phonological contrast. As such, that is, before they are associated to a phonetic correlate in an individual language, and by individual speakers upon acquisition, they lack any phonological identity, are interchangeable and hence arbitrary.[11] In fact, as seen under (2), they are arbitrary in three ways:[12]

(2)   Arbitrariness
   a.   the prime itself is arbitrary
        Calling primes alphas, betas etc. expresses their lack of phonological and phonetic identity: they are arbitrary items without cross-linguistically stable phonological identity. In individual languages, they are the vehicles of contrast: melodically relevant items are distinguished by calling them alpha, beta or gamma – but anything could do this job: a, b, c or 1, 2, 3 or ⌂, ℘, ⌐. Both the names of these distinction vehicles and their assignment to a given melodically relevant item that has a phonetic correlate is arbitrary.
   b.   the computation of primes is arbitrary
        Any segment may be turned into any other segment in any context and its reverse. That is, phonological computation may add or remove substance-free primes (or modify their values) in any way, and the trigger may be anything and its reverse.

---

[11]The structuralist heritage in this setup is obvious: the function of phonological primes is to assure contrast. It is true, though, that their second function, serving as support for the computation of phonologically active classes (natural classes in substance-laden approaches, see Mielke 2008), was not on the structuralist agenda. The elimination of phonetics from phonology also makes good on the Saussurian distinction between *Langue* and *Parole*: there is no *Parole* in *Langue*, and *Langue* is a kind of algebra of abstract units ("la langue est pour ainsi dire une algèbre" de Saussure 1972: 168). On this issue see also Berent (2018).

[12]As was mentioned in footnote 10, the Concordia version of SFP does not abide by (2c) (see section 5.2.4).

c. the association of a prime with a phonetic category is arbitrary

Any phonetic category may in principle be associated to any melodic prime: there are no cross-linguistically stable or innate associations. The symbol "↔" materializes the arbitrary character of the association just like at the upper interface, where any morpho-syntactic property may be associated to any phonological item: there is no cross-linguistically stable or innate association at this interface either.

At both interfaces, the associations implemented by a given language are language-specific, learned during L1 acquisition and stored in long-term memory: they are a list of spell-out instructions (which is subject to diachronic evolution). That is, the association PAST TENSE ↔ -ed in English is dictated by the properties of English, but is arbitrary in the sense that it may, in another language past tense, be associated to another phonological item. The same holds for spell-out instructions at the phonology–phonetics interface, for example, α ↔ labiality.

The computation of substance-free, phonologically meaningless primes is necessarily arbitrary (2b): in presence of colourless alphas, betas etc., it is not even possible to talk about "natural" or "unnatural" processes, or to distinguish any type of licit vs. illicit event for that matter. Substance-free primes do not "know" how they will eventually be pronounced and hence all logically possible processes based on them are equally probable, possible, and legal.[13] Substance-laden n → ŋ / __k is "natural" and ŋ → m / __r is not, but the same processes expressed with alphas and betas cannot even be judged for naturalness: α → β is not any more or less "natural" than α → γ since naturalness is not defined for alphas, betas and gammas. Only objects with phonetic properties can be more or less natural, but the very essence of alphas and betas is to not bear any phonetic properties in the phonology, where phonological computation occurs.

By contrast, phonologically meaningful items that occur at and above the skeleton do have an intrinsic, cross-linguistically stable phonological value and are therefore not interchangeable. As a consequence, their computation is not arbitrary: a vowel cannot be placed in a coda during syllabification, stress cannot fall on onsets etc. The following section shows that this segregation between the arbitrary workings of phonologically meaningless and the non-arbitrary workings of phonologically meaningful items is supported by an interesting empirical pattern.

---

[13]But there may be formal restrictions on what a possible process is: Reiss (2003: 230) argues that, for example, the computational system must not be able to produce a situation where two segments will have opposite feature values for all or a given subset of features. This is still expressible when features are alphas, betas etc. But it is not, if they turn out to be monovalent, as Odden (2022) argues for. Hence some decisions need to be made even with substance-free primes, which may be monovalent or binary, of the size of SPE features or bigger, holistic items such as Elements in Government Phonology (or primes in Dependency Phonology and Particle Phonology). These issues are orthogonal to the purpose of the present article.

2.3. Crazy rules are only ever melodically crazy

Below, the case of so-called crazy rules is examined, with a specific mention of crazy closed syllable lengthening or open syllable shortening, which are often subject to a misunderstanding.

2.3.1.   Survey of crazy rules

A crazy rule is one that does not make sense phonetically speaking, that is, which is not "natural". Since Bach and Harms (1972), a small literature on crazy rules has arisen that documents particular cases such as i → u / d__ (for Southern Pomoan, see Buckley 2000, 2003) (Vennemann 1972, Hyman, 2001). Chabot (2021) has drawn an inventory of cases mentioned in the literature. A relevant generalization is that crazy rules appear to only ever be melodically crazy (see also Scheer 2015: 333ff). That is, in a crazy rule A → B / C, A and B are only ever items that occur below the skeleton. There do not appear to be crazy rules that manipulate items at and above the skeleton: compensatory shortening, closed syllable lengthening (see section 2.3.2), tonic shortening etc. (syllable structure) or anti-Latin stress (stress falls on the antepenultimate syllable except when the penultimate syllable is short, in which case this syllable is stressed) are not on record.

But crazy rules not only appear to spare items at and above the skeleton: they also seem to make no reference to them in the conditioning context: i → u / d__ is reported for Southern Pomoan, but rules like p → r in coda position, or p → r after tonic vowels do not appear to occur.[14]

Of course Chabot's (2021) sample of crazy rules is incomplete and there is no guarantee that there are no syllabically or stress-wise crazy rules out there. But given the arbitrary and incomplete sample, there is no reason for all documented cases to accidentally concern only melodic properties.

Therefore the absence of crazy rules concerning syllable and stress properties is likely to be significant. This absence is exactly what is predicted by SFP: the computation of phonologically meaningless items (alphas and betas that occur below the skeleton) is arbitrary (see (2b)), but the computation of phonologically meaningful items (at and above the skeleton) is not.

2.3.2.   There is no closed syllable lengthening or open syllable shortening

The notions *closed syllable lengthening* and *open syllable shortening* are frequently misunderstood. What is meant are processes whereby the syllabic position causes

---

[14]Very rare cases of final voicing are reported, that is, languages where voiceless obstruents appear to become voiced in word-final position. Somali (Cushitic) and Lezgian (Nakh-Daghestanian) are discussed in the overview article by Iverson and Salmons (2011: 1628ff). Final voicing appears to be crazy, and its occurrence may be said to be defined by final codas: that would make a crazy rule with a syllabic conditioning. But as Iverson and Salmons report, the cases on record are disputed and alternative analyses have been proposed that make the rule non-crazy. Also consider that it may be the case that final devoicing (and hence putative final voicing) is phonetic, rather than phonological in kind (Scheer 2020).

the effect observed, just as in regular closed syllable shortening, where shortening is specifically triggered by closed syllables, and by no other factor. That is, it occurs in all closed syllables and only in this environment. A valid closed syllable lengthening (or open syllable shortening) will thus be a process whereby all short vowels lengthen specifically in closed syllables and nowhere else (or all long vowels shorten specifically in open syllables and nowhere else).

Of course, there are cases where vowel lengthening in closed syllables occurs but is not caused by syllable structure. A recurrent pattern of this kind is due to voice-induced vowel lengthening whereby short vowels lengthen before a voiced consonant, which may be a coda. This pattern is found in German, in the evolution of Western Slavic (Scheer 2017), in English (where it may be called pre-fortis clipping) and beyond (Chen 1970, Klatt 1973). Like other processes, it has a phonetic origin and may be phonologized in systems with distinctive vowel length (like German or Czech). When vowels lengthen before voiced consonants that happen to be codas, there is of course no closed syllable lengthening, since closed syllables play no role in the causality of the process.

In this context, an interesting case is reported from Menomini (Algonquian). In this language where vowel length is distinctive, Bloomfield (1939) describes vowel length alternations that respond to a complex set of conditioning factors.[15] In §32 of the article, Bloomfield reports that "if the even (second, fourth, etc.) syllable after the next preceding long vowel or after the beginning of a glottal word, is open and has a long vowel, this long vowel is replaced by short." Conversely, "if the even syllable (as [defined] in §32) is closed and contains a short vowel, this short vowel is replaced by a long" (§33). Glottal words are those "whose first syllable contains a short vowel followed by ?" (§31). That is, in order for a long vowel to shorten in an open syllable, or for a short vowel to lengthen in a closed syllable, the syllable at hand must be even-numbered with respect to the next long vowel to its left, or to the beginning of the word in glottal words. The even-numbered condition is translated into "head of a disyllabic (iambic) foot" by Hayes (1995: 219), but this is not obvious, since the calculus is not based on the left word edge (except for glottal words), but rather on the next preceding long vowel.

In any case, the Menomini process described does not qualify as closed syllable lengthening or open syllable shortening, since the causality is multi-factored: it is not the case that all long vowels in open syllables undergo shortening, or that all short vowels in closed syllables lengthen.

2.4. An arbitrary and a non-arbitrary phonology

Below, the situation described thus far is summarized and considered with regards to modularity.

---

[15]Goddard et al. (1972) report that the description in Bloomfield (1962) is incorrect, while its original statement in Bloomfield (1939) is correct. I therefore refer only to the latter.

2.4.1.   *Below and above the skeleton*

SFP thus produces a situation where an arbitrary and a non-arbitrary phonology coexist in the sense of (2a) (objects), and (2b) (the computation of these objects). Arbitrary phonology occurs below, non-arbitrary phonology at and above the skeleton in a regular autosegmental setting. This is shown under (3), which also mentions a consequence of the fact that only items below the skeleton have a phonetic correlate: these, but not items at and above the skeleton, are spelled out. This is because spelling out an item is the process of assigning it a phonetic value: items at and above the skeleton cannot be spelled out, because they do not have any phonetic correlate (more on this in section 6).

For the time being, the two phonologies are i) taken to be static (i.e., areas in the autosegmental representation, computation therein is discussed below) and ii) are called Phonology I and II. They occur in a box called production phonology under (3), referring to the computation that turns an underlying (lexical) representation into a surface representation (which is then subject to spell-out). It is opposed to lexicalization phonology, to be introduced in section 3, which includes those phonological operations that occur upon the lexicalization of morphemes, that is, when the acoustic signal is transformed into a lexically stored item. The setup under (3) will be refined as the discussion unfolds.

(3)   Two phonologies (A)[16]

----
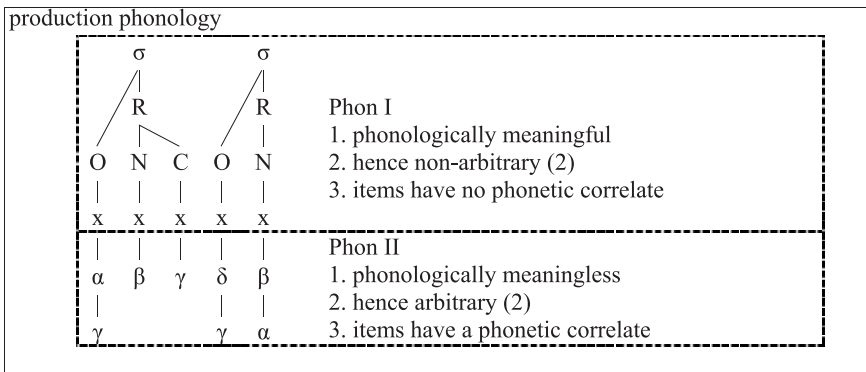
[16]Here and below, representations are used that are as theory-neutral as possible. The onset-rhyme structure is meant to represent a common denominator for autosegmental syllable structure. Nothing hinges on its properties, and alternative incarnations of syllable structure would be just as well suited to illustrate the purpose.

The coexistence of an arbitrary and a non-arbitrary phonology has an obvious consequence: we are dealing with two distinct computational systems, that is, two distinct modules. Computation within a single module could not possibly be arbitrary and non-arbitrary at the same time, or build phonologically meaningful items (onsets, feet etc.) on the basis of phonologically meaningless primes (alphas, betas, gammas).[17]

The existence of two distinct phonological modules in production raises the non-trivial questions under (4) below.

(4) Two phonologies (B)
    a. bottom-up communication
        Syllable structure is a function of items below the skeleton, that is, of the sonority of segments as well as of their linear order. How could arbitrary units whose computation is also arbitrary produce non-arbitrary syllable structure?
    b. top-down communication
        Among other phenomena, lenition and fortition that are visible on segments are a function of syllable structure. How could non-arbitrary syllable structure be trans-mitted to the arbitrary part of phonology, that is, influence arbitrary alphas and betas?

In both cases under (4), the communication between the two phonologies is non-arbitrary: syllable structure is not just anything and its reverse, given (segmental) sonority, and the lenition / fortition that is visible on segments is not random (lenition occurs in weak, not in strong positions).[18]

2.4.2.   Modular workings

The previous section has drawn the conclusion that there must be two distinct computational systems in phonology, hosting phonologically meaningless and meaningful items, respectively.

Space restrictions precluding more detailed discussion of modularity, the follow-ing merely mentions some relevant properties. Fodorian modularity (Fodor 1983) expresses the insight that computation in the mind/brain is not all-purpose, but rather falls into a number of specialized systems that are designed to carry out a spe-cific task. Computational systems are thus competent for a particular domain and can only parse information pertaining to this domain (domain specificity): their input is restricted to a specific vocabulary. That is, foreign vocabulary is unintelligible and communication with other modules requires a translational mechanism (spell-out). Humans are predisposed to develop modules for specific domains: the information

---

[17]Relevant properties of modular theory are discussed in section 2.4.2 below.

[18]Of course lenition may also occur in strong position: this is the case in spontaneous sound change, that is, when a given segment is modified in all of its occurrences regardless of pos-ition. But there is a non-arbitrary positional control on lenition and fortition: lenition may occur in strong position, but only when it also occurs in weaker positions. Cases where, say, lenition is observed word-initially (strong) but not intervocalically or in codas (weak) do not appear to be on record. That is, positional strength is relative, rather than absolute (Ségéral and Scheer 2008a: 140–143).

that there will be a computational system for vision, audition, etc. (but not for filling in a tax declaration or playing the piano) is genetically coded and present at birth. Relevant literature includes Segal (1996), Coltheart (1999), Gerrans (2002), Jackendoff (2002) and Carruthers (2006: 3ff).

Modules are thus computational systems that take a domain-specific vocabulary as an input, carry out a computation and return structure. For example, syntactic computation (Merge) operates over number, person, animacy etc. and builds hierarchical structure, the syntactic tree. In this article, vocabulary items are called primes, and the result returned by computation based on them, is structure. Hence the computation that takes sonority primes (Son) as an input returns structSon (syllable structure, feet, etc.), and the same goes for Place primes (StructPlace) as well as Lar primes (StructLar).

3. WHAT HAPPENS UPON LEXICALIZATION

Children (upon first language acquisition) and adults (when they learn new words such as acronyms, loans etc.) create new lexical entries.[19] The input of this process is the acoustic gradient waveform (first transformed into a cognitive non-linguistic form by the acoustic component of the cognitive system), and its output is a discrete symbolic (phonological) representation. This much is true for sure, and undisputed. There is a massive literature on the question of how exactly the phonetic signal is converted into lexical representations that are stored in long-term memory. For overviews see, for example, Goudbeek et al. (2005) and Boersma et al. (2003).[20]

The lexicalization process filters out linguistically irrelevant properties contained in the phonetic signal, such as pitch variations caused by male / female speakers, the emotional state of speakers, or the influence of drugs (such as alcohol) on their production.

Linguistically relevant information is transformed by the lexicalization process: the gradient non-symbolic input is stored as a sequence of discrete phonological units (segments) that decompose into melodic primes (traditionally substance-laden [±lab] etc., substance-free alphas and betas in SFP). This includes temporal order: the temporal continuum is transformed into a linear sequence of discrete phonological units (primes and segments). Linear structure in phonology is represented as a sequence of timing units (the skeleton).

Traditionally, syllable structure is absent from the lexicon: rather than being constructed upon lexicalization and stored, it is built during production phonology, that is, when the multi-morphemic string created by morpho-syntactic computation is interpreted by the phonology (underlying-to-surface). There is reason to believe

---

[19]Much of this section is owed to Faust et al. (2018).

[20]The process of reducing a gradient real-world property such as the acoustic signal to a discrete set of cognitive items is called categorization, and describes human perception (vision, audition, odour, etc., see footnote 39 on colour) in general (Harnad 2003, 2005). Non-phonological linguistic categories are in the same situation: tense is categorized time, person represents categorized speech situations, number is categorized numerosity, and so forth (Anderson 2011, vol.1: 1-10, Golston 2018).

that the syllabification algorithm runs already upon lexicalization, though, and that the lexicon is thus fully syllabified (rather than containing an ordered sequence of unsyllabified segments).[21]

Beyond the creation of a symbolic and discrete representation that includes syllable structure, there is good reason to believe that lexicalization may also impose well-formedness conditions on items that are stored. This insight is formalized as Morpheme Structure Constraints (MSC) in SPE (Chomsky and Halle 1968: 171, 382): for example, in a language like English where only clusters of rising sonority can begin a morpheme, lexical items with initial #RT clusters are prohibited, that is, will never be lexicalized. Gouskova and Becker (2016) talk about a Gatekeeper Grammar in this context (more on MSCs in section 7).

So-called Lexicon Optimization embodies the same idea. Bermúdez-Otero (2003: 29) provides the following formulation (after Hale 1973: 420): prefer inputs that are well-formed outputs. Thus an item that will never be able to appear as such on the surface (e.g., an initial #RT cluster in English) will not be stored. Relevant literature includes Prince and Smolensky (2004: Section 9.3), Yip (1996), Bermúdez-Otero (1999: 124) and Inkelas (1995).

This is also what motivated Kiparsky's (1968–1973: 14ff) Alternation Condition, which was designed to rule out so-called absolute neutralization, that is, a situation where a morpheme has an underlying sequence that may never be inspected on the surface because it is modified by computation in all occurrences (such as *nightingale* whose first vowel was said to be /i/, rather than /aj/, on the analysis of Chomsky and Halle 1968: 234).

Lexicalization falls into two distinct actions according to the input. The phonetic signal is a non-cognitive, gradient (real world) item that is converted into cognitive, discrete, symbolic and linearized units: the segments (or phonemes). These segments are then the input to further computation which builds phonological structure (syllable structure, feet or equivalent stress-encoding structure in lexical stress languages).[22] Items involved in both steps are subject to well-formedness conditions. The former may be called lexicalization conversion and the latter lexicalization phonology.

The two systems are different in kind because they work on dramatically different input vocabulary: lexicalization conversion counts among the human perception

---

[21]Which of course does not preclude the existence of floating items in lexical representations. Vaux (2003) and Vaux and Samuels (2018) provide a documented overview of the various voices that have advocated both positions, that is, syllable structure being absent vs. present in the lexicon. They conclude that it is present. This is also the position taken by Government Phonology since its inception (Kaye et al. 1990, Scheer and Kula 2018).

Surely the most obvious and trivial argument is length. Since the advent of autosegmental representations, it is consensual that simple and geminate consonants, as well as long and short vowels, are not distinguished by a melodic feature. Rather, their difference is encoded by timing units (skeletal slots or morae). Since length is an idiosyncratic and contrastive property of segments, the skeletal or moraic structure at hand must be stored in the lexicon.
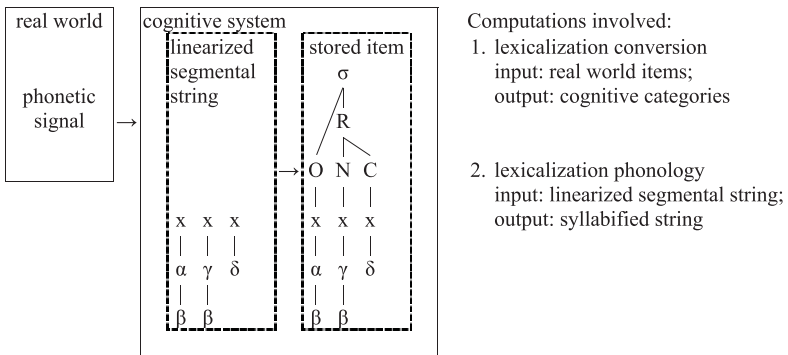
[22]In Nasukawa's Precedence-free Phonology (Nasukawa 2015, Backley and Nasukawa 2020), lexicalization phonology combines primes (Elements) into larger units of segment and syllable size.

devices that transform real-world items into cognitive categories: light waves into colours, chemical substances into odors etc. (see footnotes 20 and 39). By contrast, lexicalization phonology is just like production phonology (see (3)) that occurs when the multi-morphemic string created by morpho-syntactic computation is interpreted (underlying-to-surface): it takes as an input cognitive, discrete and symbolic units that belong to the phonological vocabulary (primes); it produces phonological structure and assesses well-formedness.[23]

Obviously there is some overlap between the phonological operations that are carried out in production and upon lexicalization: syllabification, for example, is active in both. There is reason to believe that they are not identical, though. This issue, known as duplication and regarded as a problem by some, is further discussed in the conclusion (section 7).

Summing up the discussion, (5) below depicts what happens upon lexicalization (allowing for the shortcut between the acoustic real-world object and the linguistic system: as was mentioned, cognitive processing in the acoustic component occurs between the two boxes shown).

(5)   Two computations upon lexicalization

Computations involved:
1. lexicalization conversion
   input: real world items;
   output: cognitive categories

2. lexicalization phonology
   input: linearized segmental string;
   output: syllabified string

---

[23]The overall architecture pieced together from lexicalization (this section) and production (sections 2.4, 5) is akin to the BiPhon model (Boersma 1998, Boersma and Hamann 2008) in a number of ways (including substance-freeness). The heart of BiPhon is to assume that the same computational system is responsible for the operations in perception and production (hence the name of the approach). That is, gradience and discreteness are converted into one another by the same computational device in both directions.

Lexicalization conversion is perception but is not necessarily identical to the perception that occurs when speakers work on the acoustic signal in order to access its meaning in a verbal exchange: nothing is lexicalized here. Therefore, lexicalization is the word used in this article for the process that leads to storage in long-term memory. Perception is more general, and its devices likely overlap with lexicalization (see section 7). Also, linearization upon lexicalization (of segments within a morpheme) and in production (of morphemes) appears to be quite different in kind.

4. SONORITY IS DIFFERENT

Section 3 has described how the items that are used in production come into being, that is to say, enter the lexicon. They have the structure discussed in section 2, that is, what was called Phon I and Phon II under (3). For the time being this difference is only grossly depicted as two areas in an autosegmental representation. The present and the following sections take a closer look at how Phon I (at and above the skeleton) comes into being, what kind of items do and do not appear in this area, as well as how Phon I and Phon II communicate.

Below, the traditional partition of segmental properties (as embodied in Feature Geometry, Clements and Hume 1995) is referred to as Son (sonority defining primes), Place (place defining primes) and Lar (primes defining laryngeal properties).[24]

4.1. Diagnostics

Below, a number of diagnostics are discussed which suggest that sonority is different.

4.1.1.   To be heard vs. to be understood

Harris (2006) argues that Son is different in kind with respect to Place and Lar. Observing that "sonority differences are never contrastive in the way that differences defined in terms of individual features can be", he argues that phonetically speaking, "speech is a modulated carrier signal: the modulations bear the linguistic message, while the carrier enables the message to be heard." That is, the carrier signal, sonority, "is concerned primarily with the audibility of the linguistic message rather than with the message itself" (all quotes from Harris 2006: 1484). Thus sonority is about being heard, while place and laryngeal properties are about being understood: they carry the linguistic message that sonority is not conveying.[25]

4.1.2.   Selective bottom-up visibility

It is a trivial, though rarely explicitly stated fact that Son is projected above the skeleton, while Place and Lar are not. Syllable structure is a function of two and only two factors: the linear order of segments and their relative sonority. That syllable structure depends on these two factors (plus parametric settings determining whether or not specific syllabic configurations such as codas, branching onsets etc. are provided for) and on no other is an undisputed and theory-independent fact which is transcribed in all syllabification algorithms (e.g., Steriade 1982: 72ff, Blevins 1995: 221ff, Hayes 2009: 251ff).

---

[24]Sections 4.1 and 4.2 are an updated, digest version of Scheer (2019b).

[25]From this difference, Harris (2006) concludes that sonority lies outside of grammar. He does not dwell on the consequences of this claim, though: if there is no sonority in grammar, there is no syllable structure. What, then, is the massive body of syllable-based phenomena about? And, more down to earth, how would grammar be able to make a difference between, say, vowels and consonants in trivial processes such as vowel harmony? These issues are further discussed in section 5.2.3.

That Son, but not Place and Lar, are present above the skeleton is also shown by reverse engineering. Sonority values may to a certain extent be predicted given syllable structure: onsets and codas contain low sonority items, nuclei contain high sonority items; the sonority of the members of a branching onset is rising; it is non-rising in a coda-onset sequence. By contrast, syllable structure does not allow us to deduce anything about Place or Lar properties of segments: the fact of being a nucleus, onset or coda tells you nothing about whether the segments associated are labial, palatal, velar etc., and onset- and codahood provide no clue whether segments are voiced or voiceless either.[26]

There is thus reason to believe that Phon I under (3), that is, items which occur at and above the skeleton, are the result of a structure-building computation that is based on (phonologically meaningful) sonority primes. That is, the domain-specific vocabulary of the Son module are Son primes, and modular computation based on them returns structSon, the items familiar from above the skeleton.

4.1.3.   *Selective top-down visibility*

The previous section has discussed the fact that syllable structure is built without reference to Place and Lar. This indicates that the Son vocabulary on the one hand, and Place / Lar vocabulary on the other hand, are distinct: if Place / Lar primes were present and available when structSon is built, they would be expected to be taken into account. The fact that they are not suggests that they occur in a different module.

This conclusion is supported by a solid empirical generalization: processes which occur at and above the skeleton may be conditioned by Son, but are blind to Place and Lar.[27] Evidence comes from linearization when the landing site of infixes is defined (see (6a)) from phonologically conditioned allomorph selection (6b), from stress placement (6c), from contour tone placement (6d) and from positional strength (6e). If the processes at hand occur in the Son module that is different from the Place / Lar module(s), it follows that they cannot see Place / Lar properties (just as syllabification is blind to them).

---

[26]Except of course in case a relevant process or restriction is active in individual languages: coda consonants will be voiceless in a language that has final devoicing, and if a language restricts coda consonants to coronals or velars, Place may be predicted from codahood. But this is orthogonal to the point made: the predictability at hand is due to the existence of a language-specific process or coda restriction, not to codahood as such. In contrast, a branching onset allows for the deduction of the sonority of its members as such, universally and regardless of language-specific settings.

[27]This generalization is a piece of a broader project, melody-free syntax (and morphology), whose goal is to show that morpho-syntactic computation and items below the skeleton, or rather: Place and Lar, are incommunicado in both directions. Melody-free syntax is developed in Scheer (2011: Sections 412, 660, 2019b); also in a number of conference presentations since Scheer (2012b), and in Scheer (2016) regarding phonologically conditioned allomorphy.

(6) Selective visibility from above
    a. Linearization (infixation)
        The list of anchor points that infixes look at in order to determine their landing site falls into two categories: edge-oriented and prominence-oriented. For the left edge, for example, documented situations are "after the first consonant (or consonant cluster)", "after the first vowel", "after the first syllable" and "after the second consonant". Prominence-based attractors are stressed vowels, stressed syllables, or stressed feet.

        In no case are Place or Lar reported to be relevant for the definition of the landing site. Hence cases where infixes are inserted after, say, the first labial or voiceless consonant in the word (and in their absence are prefixed) do not appear to be on record. This typological generalization is made by Moravcsik (2000) and Yu (2007). The latter studied 154 infixation patterns in 111 languages belonging to 26 different phyla and isolates. See also Samuels (2009: 147ff).

    b. Phonologically conditioned allomorphy
        The selection of allomorphs, that is, independently stored variants of the same morpheme, may be conditioned by purely morpho-syntactic factors (like in the case of *go* and *went*, selected according to tense), but phonological properties may also play a role (as in Moroccan Arabic where the 3sg masculine object/possessor clitic is *-h* after V-final, but *-u* after C-final stems). Phonological factors boil down to Son, though: conditioning factors regarding Place and Lar are not on record.

        This is a typological generalization established in Scheer (2016) based on the empirical record of Paster (2006), who has surveyed about 600 languages and described 137 patterns in 67 languages. Seven cases that may appear problematic are analyzed in Scheer (2016).

    c. Stress
        Stress placement is known to be conditioned by syllable structure: so-called Weight-by-Position regulates whether or not closed syllables count as heavy. In some languages a more fine-grained distinction is found, though, appealing to Son: sonorant, but not obstruent codas, contribute to the weight of their syllable (e.g., Wilson 1986, Zec 1995: 103ff, Szigetvári and Scheer 2005: 44ff; see also the typological survey in Gordon 2006). But there is no case on record where Place or Lar has this effect ("a coda is heavy only if it is labial / voiceless").

        On the vocalic side, de Lacy (2002) and Gordon (2006: 52) have established the same generalization: in many languages, stress placement is sensitive to the sonority of vowels (low, mid, high), but Place and Lar never play a role (de Lacy 2002: 93).[28]

    d. Tone
        The situation for tone is much the same as for stress. In many languages, contour tones may only appear on heavy syllables, and what counts as heavy may be determined by the presence or absence of a long vowel and a coda, as well as by coda sonority. Place and Lar are never involved in the definition of what counts as light or heavy, though.

        This generalization is based on the typological work by Gordon (2006: 34, 85), who has studied some 400 languages.

---

[28]Note that the status of vowel height may be subject to debate: being classically understood as a Place property, it is nonetheless the representative of sonority in the vocalic realm.

e.   Positional strength of post-coda consonants
     Positional strength, that is, lenition and fortition, is a function of syllable structure
     (initial and post-coda position is strong, intervocalic and coda position is weak),
     except for a cross-linguistic parameter that controls the behaviour of consonants in
     post-coda position. In this context, consonants may either be strong no matter what,
     or only after obstruents (while following a weak pattern after sonorants). The reverse
     distribution (i.e., strength after sonorants, weakness after obstruents) does not appear
     to exist. Also, Place and Lar have no bearing on post-coda strength (there is nothing
     like "post-coda is weak if the coda hosts a labial / voiceless consonant").
     This pattern is discussed by Ségéral and Scheer (2008a: 156–161).

Note that the diagnostic for tone provided by contour tone placement (6d) appears to be consistent with the fact that tone may also influence allomorph selection (6b): Paster (2006: 126–130) discusses a case from the Yucunany dialect of Mixtepec Mixtec (Otomanguean, Mexico) where the first person sg is marked by a floating low tone, except when the final tone of the verb stem is low, in which case *-yù* is suffixed. This also ties in with the fact that tone appears to be able to influence morpho-syntactic computation, as reported by Rose and Jenks (2011) (see section 5.4.2). These three diagnostics are consistent, suggesting that tone belongs to the area above the skeleton (the Son module) where it is typically located in autosegmental representations. But they conflict with the fact that tone has a phonetic correlate (pitch) and interacts with regular melodic primes (namely in Lar) (see footnote 8). The diagnostics developed in the present article thus confirm the infamous hermaphrodite nature of tone (Hyman 2011).

4.2.   The representation of sonority

The preceding sections have provided evidence that Son is different in kind with respect to Place / Lar. There are two lines of thought that share the same diagnostic, albeit on entirely different grounds, and concluding that there are no Son primes (i.e., phonological objects that specifically represent sonority). In one, sonority is phonological but a property of structure, rather than of primes; in the other, syllable structure is the result of phonetic perceptibility, rather than of a syllabification algorithm that runs in the phonology and is based on sonority primes. Both approaches are discussed below, starting with the latter.

4.2.1.   *Perception-based sonority*

Perception-based sonority is exposed by Steriade (1994, 1999). As Moreton et al. (2005: 341) put it, "perceptibility theory […] says that a segment's compatibility with a given environment depends on how accurately it is likely to be perceived in that environment", where "environment" refers to classical syllabic positions such as onset, coda and so forth. In this approach, syllable structure exists in phonological representations, but is directly abstracted from phonetic information and has no bearing on segmental alternations. That is, VTRV is parsed as V.TRV in Spanish but as VT.RV in Classical Arabic, but "consonantal phonotactics are best understood

as syllable-independent, string-based conditions reflecting positional differences in the perceptibility of contrasts" (Steriade 1999: 205). In this view, lenition and fortition are thus entirely unrelated to syllable structure, the latter however existing independently.

The question whether syllable structure is responsible for positional phenomena is unrelated to the issue pursued here, that is, the existence of Son primes and the way items, at and above the skeleton, are built. Regarding this issue, it seems to me that the perceptibility-based approach is in fact a notational variant of the regular perspective based on Son primes. When building (phonological) syllable structure based on perceptibility, the gradient phonetic signal will also need to be transformed into discrete phonological units, the segments, and it must also be somehow decided that certain segments are vowels, while others are consonants, and that the former will end up in a nucleus, while the latter are parsed into an onset or a coda. This is but a short version of the process described in section 3 and further discussed in section 5.2.3, that is, the construction of syllable structure based on the phonetic signal with an intermediate classification of segments in terms of sonority.

4.2.2.   *Structuralization of sonority*

In SPE, major class features such as [±son], [±cons] or [±voc] were scrambled in a single feature matrix together with features defining melodic characteristics such as place of articulation and laryngeal properties. OT implements sonority along the same lines, except that scrambling concerns constraints, rather than features (e.g., Smith and Moreton 2012).

Departing from an amorphous and unstructured set of primes, Feature Geometry made a first step towards the structural representation of sonority. The theory introduced by Clements (1985) autosegmentalized the amorphous bundle of SPE features, aiming at grouping them into natural classes (class nodes) in a feature geometric tree. In this approach, sonority is still represented in terms of the SPE features (or some version thereof), but following the class node logic, features are grouped together and isolated from other classes of features. In Clements and Hume's (1995) final model of Feature Geometry, the Son primes [±son], [±approximant] and [±vocoid] are borne directly by the root node, as opposed to other features which are represented in the feature geometric tree below the root node.

This structural insulation of sonority was further developed in a number of approaches sharing the idea that rather than being represented by specific primes, sonority is a function of Place and Lar primes as well as of their organization. Government Phonology has pioneered this line of thought: in this approach, sonority is a derived notion based on segmental complexity. That is, the more primes a segment is made of, the less sonorous it is (and vice versa) (Harris 1990, Backley 2011: 114ff). Segmental complexity has a number of issues: for example, it cannot derive vocalic sonority. High *i,u* and low *a* are made of just one prime (I, U and A, respectively), while mid vowels bear two primes (I,A or U,A) and should thus be less sonorous than high and low vowels, which of course is not the case.

Other implementations of the idea that sonority is a structural property include GP2.0 (Pöchtrager 2006: 55ff, Pöchtrager and Kaye 2013); Rice (1992); Schwartz' (2013, 2017) Onset Prominence model; Hulst's (1994, 1995a, 1999) Radical CV Phonology; and de Carvalho's (2002a, 2008, 2017) perspective. These approaches are reviewed in greater detail in Hermans and Oostendorp (2005) and in Scheer (2019b).

The visibility-based evidence discussed in sections 4.1.2 and 4.1.3 documents the exclusive reference to sonority for the processes mentioned, with Place and Lar being irrelevant and invisible. Given this situation, it would come as a surprise if sonority and melody cohabited the same space: were they all accessible, there would be no reason why some items should be selectively taken into account, while others are actively ignored. Therefore, Place and Lar information must not be accessible when sonority computation is carried out, irrespective of how sonority is represented (structuralized or in the guise of primes). Scheer (2019b) thus concludes that the empirical situation regarding visibility warrants a segregation of sonority and Place / Laryngeal properties into distinct modules, rather than the structuralization of the former.[29]

Approaches that make sonority a structural property thus take a step in the right direction by estranging sonority from Place / Lar. But they leave all players in the same computational space (module). If a further step is taken that recognizes the segregation of Son and Place / Lar in distinct computational systems, the structural representation of sonority loses its *raison d'être*, which is to enact the difference.

If thus sonority is a computational system of its own, there must be Son primes, since modular computation is always based on primes (see section 2.4.2).

4.3.   The Place and Lar modules

Place and Lar computation appear to be mutually waterproof. While of course Place can impact Place (e.g., velar palatalization) and Lar bears on Lar (e.g., voicing assimilation), Place modifications triggered by Lar (a velar is turned into a palatal before a voiceless consonant) or the reverse (a voiceless consonant becomes voiced when followed by a labial) appear entirely outlandish and stand a good chance of being absent from the record.[30]

---

[29]A reviewer raises the issue of whether certain groups of syntactic features could or should also be segregated into distinct modules. The question is well taken, but falls beyond the scope of the article (and of my own competence). Adger and Svenonius (2012) identify different classes of syntactic features, but not on the grounds of their mutual visibility: rather, a feature class in syntax is one where the members share some syntactically relevant property. For example, agreement concerns [person], [number] and [gender] but not, say, [past] or [present]. Following this classification, features are not in complementary distribution regarding syntactic properties, though: for example, C and v are phase heads (in Chomsky's 2000 system), but v and V are θ-assigners.

[30]Lombardi (2001) concludes that Place and Voice are different because the cross-linguistic reaction against coda consonants is in complementary distribution according to whether the coda offense is due to place or to voice. While illegal laryngeal configurations in codas are resolved by neutralization (i.e., the shift to a legal laryngeal value), illegal place patterns in codas provoke epenthesis of a vowel or deletion. This reaction is never found for laryngeal

In account of their difference, Feature Geometry has devised distinct category nodes for Place and Lar. In the same way as for the difference between Son vs. Place / Lar discussed in the previous section, structuralizing the distinction between Place and Lar is a step in the right direction, but does not forbid Place to condition a Lar process, or the reverse, since both coexist in the same computational space. Locating Place and Lar in different modules enacts the fact that process where one conditions the other appear to be absent from the record: they are excluded, since each computational system can only bear on items of its own domain; see under (7).

4.4. Impact of structSon on Son, Place and Lar

Section 4.1.3 has shown that there is no bottom-up conditioning: Son (at and above the skeleton) is not influenced by Place or Lar (which live below the skeleton). Top-down conditioning (where syllable structure or stress condition Place and Lar) exists, though.

(7) Top-down conditioning
    a. structSon bears on Lar
        1. devoicing (or more generally delaryngealization) in coda position.
        2. lenition
            typical syllable-conditioned trajectories along the sonority scale involve laryngeal stages, for example, $p^h > p > b > v$ in intervocalic position.
    b. structSon bears on Place
        1. lenition: debuccalization $f > h$, $s > h$ in coda position.
        2. unstressed vowels centralize (become schwa).

These patterns are unexpected if Son and Place / Lar are located in different modules. But consider the fact that regular Place and regular Lar computation appears to be unimpacted by structSon: there is nothing like "velars palatalize before front vowels, but only if they belong to a coda" or "obstruents undergo progressive voice assimilation, but only if they are engaged in a branching onset".

This suggests that Place- and Lar-internal computation is indeed insensitive to structSon, as predicted. The patterns under (7) are different: rather than bearing on regular Place- and Lar-computation that exists by itself, they describe a situation where a Place or Lar modification may occur when *triggered* by structSon. How could that be possible in a modular environment?

In order to approach this question, let us take a closer look at lenition and fortition. Inoffensive cases modify only Son primes, given structSon: for example, a segment whose Son primes define a stop is turned into a fricative with corresponding Son primes because it occurs in intervocalic position. The input, conditioning context and output of this process are all located in Son and that complies with modular standards.

---

coda offenses, and neutralization (e.g., an illegal velar becoming dental) is absent from the record for place-based coda offenders.

But as was mentioned, there are also offensive cases: typical lenition trajectories move along the sonority scale but also involve steps that are defined by Lar ($p^h > p > b > v$ etc.) (7a2). This is reflected in sonority scales that scramble Son and Lar properties (Szigetvári 2008a, Parker 2011: 1177 counts 17 steps). The type of lenition mentioned under (7b1), debuccalization, involves the loss of Place primes and may thus not be described as a rearrangement of Son primes either. Szigetvári (2008b) shows that these two types of lenition exhaust the empirical record: one makes segments more sonorous (thus involving Son primes: $t > \theta$, $b > \beta$, $t > \mathrm{f}$ etc.), the other makes them lose Place primes, as in $t > \mathrm{?}$, $s > h$, $f > h$.

In Government Phonology, lenition (of both types) is interpreted as a loss of (privative) primes (Harris 1990). Although structSon cannot see or bear on Place primes across a module boundary, it may influence the properties of timing units. Following the GP logic, timing units that are weak because their associated syllabic constituent is weak are unable to sustain other primes in other modules. That is, positions, rather than segments, are weak or strong: this is quite trivial and encoded in the name of the phenomenon, *positional* strength. The effect that is visible on segments is the consequence of their positional strength or weakness, that is, of the positional status of their x-slot.

This mechanism may be generalized: structSon bears on timing units and makes them more or less apt at sustaining (licensing) primes that are associated to them. A strong position is a position that allows for large number of primes to be associated, while a weak position can sustain only a few, or indeed no prime at all. Timing units do not belong to any individual module, but rather instantiate the temporal and linear sequence. The primes of all three modules that co-define a segment are attached to the timing unit that defines this segment (more on this in sections 5.5.1 and 5.5.2, see the representation under (11) in section 5.3). Hence restrictions that weigh on a timing unit will be visible in all three modules.

All cases under (7) share the fact that under the pressure of structSon some prime is lost: a Lar prime under (7a1) (devoicing); Son primes ($b > v$) or Lar primes ($p > b$) under (7a2) (sonority-lenition); Place primes under (7) (debuccalization-lenition); and (7b2) (centralization of unstressed vowels). Hence the analysis whereby structSon weighs on timing units whose weakness (or strength) then has consequences on Place and Lar is workable throughout.

A structSon conditioning on timing units may lead to the loss of a segment in diachronic lenition (loss of a coda consonant in $s > h > \emptyset$), but also in synchronic computation: vowel-zero alternations are governed by a (cross-linguistically stable, see Scheer 2004: Section16) syllabic context where lexically specified vowels (schwas, short vowels, high vowels etc.) are absent in open, but present in closed syllables. In Czech for example, the root vowel of *pes* "dog Nsg" alternates with zero in presence of the case marker *-a* in *ps-a* "dog Gsg" (while the phonetically identical root vowel in *les - les-a* "forest Nsg, Gsg" is stable: alternating *e* is marked as such lexically).

In Strict CV, alternating vowels are underlyingly present but floating: the vowel under (8a) fails to be associated when its nucleus is governed. This is the case in Gsg where the following nucleus hosts *-a* (8c). In Nsg under (8b), though, the following

nucleus is empty and hence cannot govern. Therefore, the nucleus containing the alternating vowel remains ungoverned, which allows the floating melody to associate.

(8) Czech *pes* - *ps-a* "dog Nsg, Gsg"
    a. lexical entry   b. Nsg pes   c. Gsg ps-a
                                  gvt   gvt

```
    C  V  C  V        C  V  C  V         C  V  C  V
    |     |           |  ↑  |           |     |  |
    p  e  s           p  e  s           p  e  s  a
```

Note that in Strict CV, government (and licensing) represent syllable structure, that is, structSon. But whatever the framework used, the presence or absence of the alternating vowel is controlled by syllable structure. In Strict CV, its absence is due to the pressure that structSon puts on the nucleus (government). This is parallel to what we know from lenition: the ability of a constituent to sustain primes is curtailed under positional pressure. In the case of vowel-zero alternations, its ability to host primes is zero.

In sum, structSon may indirectly cause the loss of primes in all three modules by defining strong and weak timing units, which then may not be able to tolerate the association of certain primes. But there is no bearing of structSon on regular computation that occurs in the Place or the Lar module.[31] This setup may thus be said to kill two birds with one stone.

5. WORKINGS OF PHONOLOGY WITH THREE MODULES

In this section, the ingredients of a full-blown SFP are considered: the nature of lexical entries and primes, computation (upon lexicalization and in production), multi-module spell-out, and phonetics (language-specific as well as universal).

5.1. Lexical entries

If Son vs. Place vs. Lar are three distinct computational systems, according to modular standards each one carries out a specific computation based on a proprietary vocabulary. A segment thus combines items from three distinct module-specific vocabularies. The situation where a single lexical entry is defined by items that

---

[31]The cross-linguistically pervasive intimacy and interaction between voicing and nasality (Nasukawa 2005) has a phonetic basis (Solé et al. 2008). The typical phonological process is post-nasal voicing, but post-nasal devoicing also exists (Solé et al. 2010). This situation suggests that nasality is located in Lar, rather than in Son.

belong to different vocabularies is known from the morpho-syntax – phonology interface where idiosyncratic properties of single lexical entries (morphemes) are used by multiple computational systems. The lexical entry of a morpheme stores three types of information written in three distinct vocabularies that are accessed by three different computational systems: morpho-syntax (using vocabulary items such as number, person, animacy, etc.), semantics (LF-relevant properties) and phonology (vocabulary items such as occlusion, labiality, voicing etc.). The lexical entry for *cat* for example may look like (9a).[32]

(9)   lexical entries
      a.   of a morpheme: *cat*
           M307
           [<Morph-synt = animate, count etc.>, <Sem = "idea of cat">, <Phon = /kæt/>]$_{\text{morph.}}$
      b.   of segments:
           P12 [<Son = stop>, <Place = velar>, <Lar = voiceless>]$_{\text{segment}}$ (= k)
           P16 [<Son = stop>, <Place = dental>, <Lar = voiceless>]$_{\text{segment}}$ (= t)
           P19 [<Son = vowel>, <Place = front low>, <Lar = – >]$_{\text{segment}}$ (= æ)

      In the same way, the lexical identity of a segment contains several compartments that host items from three distinct vocabularies, as shown under (9b) for the constituent segments of *cat*. Each vocabulary is then accessed by the relevant computational system. For instance, the computation that builds syllable structure is specific to Son and can only parse this vocabulary. Hence only Son primes are projected and the result of the computation, structSon, contains no information pertaining to Place or Lar. This is why other computations which occur above the skeleton and take into account items in that area are unable to access anything else than Son primes and structSon (section 4.1.3). In the same way, Place computation (e.g., palatalization) accesses only Place primes (and outputs structPlace), while Lar computation (e.g., voice assimilation) reads only Lar primes (and produces structLar).

5.2.   Lexicalization conversion

This section discusses the fundamental contrast between Place and Lar primes on the one hand, and Son primes on the other: it is argued that the latter are universally associated to a phonetic category, while the former are not.

5.2.1.   *Primes are module-specific*

It was mentioned in section 2.4.2 that modules are hard wired, that is, genetically coded: humans are predisposed for processing visual, auditory etc. information. If Son, Place and Lar are three distinct modules, their existence and domain of competence must thus be genetically coded and present at birth. That is, children know

---

[32]The content of Phon under (9a) is simplified: in addition to the three segments shown, linearity (the skeleton) and syllable structure are represented (see section 3).

beforehand that these three modules exist, and that their working requires the construction of relevant primes in each domain.[33]

An argument for the existence of three distinct and universal systems is the trivial fact that all languages appear to have distinctions in all three areas: there is no language which chooses, say, to ignore the possibility of implementing Lar distinctions, or which builds a distinctive system without some contrast in Place, or where all segments belong to the same sonority class. If primes were entirely unconstrained at birth, some language could build a contrastive system where, say, laryngeal properties play no role. What is observed, though, is that while the three types of contrast are always present independently of the language-specific environment, there is variation within each system.

A consequence of there being three distinct modules is that alphas, betas and gammas are not colourless, but rather specific to the three domains at hand. That is, primes occur in a specific module and are thus different in kind. This difference may be graphically indicated by using symbols from different alphabets, say, α, β, γ for Place primes, б, г, д for Lar primes and a, b, c for Son primes.[34]

5.2.2. *Place and Lar primes*

Place and Lar primes bring together the three types of arbitrariness under (2): the primes themselves are arbitrary ((2a): α is not any more appropriate than β); their computation is arbitrary ((2b): any prime may be added to or removed from any segment, given any trigger and its reverse); and so is their association to a phonetic category ((2c): any association of any prime with anyphonetic category is possible).

---

[33]The language faculty being unspecific for the vocal or signed (or still other) modalities, the domain of competence of component modules should also be modality-unspecific. That is, Son, Place and Lar may be the instantiations of the three modules at hand in the vocal modality, while movement, handshape and location are their representatives in sign language. On this equivalence see Hulst (1995b) and Sandler (1993), the latter author calling on Stokoe (1978) "in positing three major phonological categories: Hand Configuration (HC), Location (L) and Movement (M)" (Sandler 1993: 245). The fact that the analysis of both vocal and sign language has produced a distinction of three basic types of properties is likely not accidental: it is compatible with a common origin. See Sandler and Lillo-Martin (2006: 111-245) for an overview of the question.

[34]The granularity of modules, that is, the size of their domain, is an empirical question and, like objects of study in other areas (physics, chemistry, biology), has evolved from bigger to smaller items as inquiry proceeded. Vision for example was thought of as a single faculty, but today falls into a number of distinct computational systems that are specialized in shape, colour, motion, form, face recognition, or contrast (owing to much work since Marr 1982; see the overview in Stevens 2012). In the same way, Fodor (1983) talked about language as a single computational system, but it is fairly consensual today (for those who abide by modular distinctions) that language falls into distinct computational systems specializing in morpho-syntax (or morphology and syntax), semantics, phonology and phonetics (and maybe others) (see Scheer 2011: Section 622 for a historical overview). Identifying distinct computational systems within phonology is taking a further step on the path from bigger to smaller objects of inquiry.

Primes themselves are thus absent at birth. They are created and associated to relevant phonetic correlates upon exposure to contrast and phonological processing of the target language (e.g., Boersma 1998: 461ff, Mielke 2008, Dresher 2014, 2018, Odden 2022; see also Goudbeek et al. 2005).

5.2.3.   *Son primes are hard wired*

Section 2 has concluded that unlike other segmental properties, sonority is phonologically meaningful. This describes the fact that the Son module is not arbitrary. Unlike what is observed for Place and Lar, the computation of Son primes upon lexicalization is lawful (hence at odds with (2b):its result, structSon, is the same in all languages. That is, its items and their properties as well as their arrangement are universal: the list of structSon items is the same in all languages (onset, nuclei, feet, etc.), their properties are shared across languages (onsets host consonants, nuclei host vowels, or high sonority items) and their arrangement is also the same everywhere (onsets occur to the left, not to the right of rhymes, etc.).[35]

The association of a Son prime with a phonetic correlate is also not arbitrary (against (2c)). As Clements (2009: 165ff) points out, there are no flip-flop systems where all items that are vowels phonetically-speaking would be interpreted as consonants in the cognitive system upon lexicalization, and thus associate to a C position, while all items that are phonetic consonants would become phonological vowels and therefore associate to a V position. This means that Son primes are extracted from the phonetic signal in a predictable way that is at least partly the same in all languages. Clements (1990: 291) says that "in the absence of a consistent, physical basis for characterizing sonority in language-independent terms, we are unable to explain the nearly identical nature of sonority constraints across languages" (see also Clements 2009: 166).

There must thus be a way for humans to recognize the difference between consonants and vowels in the acoustic signal, and this capacity is put to use in such a way that there is a hard-wired relationship between a phonetic vowel and its lexicalization as a cognitive / phonological vowel, as well as between a phonetic consonant and its lexicalization as a cognitive / phonological consonant.

Therefore, children must roughly know what a vowel sounds like at birth, and in which way a vowel is distinct from a consonant. In an SFP setting, this means that they must have Son primes that are associated with a specific phonetic correlate. The literature has identified perceptual salience as a cross-linguistically robust correlate of sonority, which owes to phonetic properties such as loudness, energy and intensity (Bloch and Trager 1942: 22; Heffner 1950: 74; Fletcher 1972: 82ff; Price 1980; Ohala 1992; Wright 2004: 39ff; Harris 2006; Clements 2009; Parker 2008, 2011; Gordon et al. 2012; Bakst and Katz 2014).[36]

---

[35]Notwithstanding relevant parameterization, of course: some languages have branching onsets, others do not, some have codas, others do not, etc.

[36]The obvious articulatory correlate of sonority is aperture, that is, the physical distance between the upper and the lower articulator (Keating 1983, Beckman et al. 1992). The idea

Despite this massive literature, there are also voices saying that no phonetic correlate was ever identified: Hooper (1976: 198), Ohala and Kawasaki (1984: 122), Ohala (1990: 160),. Later on, though, Ohala (1992: 325) arrives at about the same conclusion as the literature mentioned: sonority is the result of a number of acoustic parameters, that is, amplitude, periodicity, spectral shape and fundamental frequency.

The idea that no phonetic correlate of sonority is currently available may either lead to the rejection of sonority as an unscientific concept, or to consider the issue an open question whose answer is orthogonal to phonological analysis. The former position was taken by the early Ohala (1974: 252) who says that phonologists "invent meaningless labels for types of sounds and sound patterns (e.g., marked, strong, sonority, chromatic)" (on this, see Hankamer and Aissen 1974: 143, footnote 8).

The latter position is adopted by Hankamer and Aissen (1974: 137), Hooper (1976: 198) and Clements (1990: 291): we know that sonority exists, since the analysis of phonological patterns shows it does, and there must be some phonetic correlate which, in the current state of our knowledge, we are not able to identify. Hooper (1976: 198) argues that whether or not analysts are able to identify the phonetic correlate of sonority by machine measurements is irrelevant for linguistic analysis. We know that the human abstracts sonority from the acoustic signal, just as we know that the human is categorizing the gradient acoustic signal into discrete segments (or phonemes), while it is impossible to identify where exactly one segment ends and another begins, based on machine or other analysis of the signal. It would of course be enjoyable, Hooper says, to know how exactly the human extracts sonority and segments from the gradient acoustic signal, but our current ignorance is not an obstacle to taking sonority or the segment as a linguistic fact.

The situation described leaves no serious doubt that sonority does have a cross-linguistically stable phonetic correlate, and that this correlate is perceptual salience (in its various acoustic guises).

5.2.4. *Sonority is phonetically composite and entropy-driven*

The literature mentioned contains two ideas that are worth isolating. The observation that sonority is phonetically composite is made, among others, by Ohala (1992: 325) and Harris (2006: 1486), the latter concluding "that sonority does not map to any unitary physical property but is rather a cover term for a collection of independent acoustic properties that contribute to an overall dimension of perceptibility or auditory-perceptual salience." The other idea is that modulations (modifications of the signal) count, rather than static values (Ohala 1992: 325, Harris 2006).

Stilp and Kluender's (2010) study is based on these two insights. It takes the latter idea to follow from Shannon's (1948) more general information theory which is based on the idea that the degree of informativeness of an event depends on its entropy, that is, the uncertainty or unpredictability associated with it: the more unexpected, the more informative. Stilp and Kluender (2010: 12387) thus

---

that sonority is based on a more or less severe obstacle produced by the articulatory system is implemented by Hume and Odden (1996) under the heading of impedance.

test "whether relative informativeness of portions of the speech signal is related to the degree to which the signal changes as a function of time." On these grounds, they develop the notion of Cochlea-scaled Spectral Entropy (CSE), which "is a measure of the relative (un)predictability of signals that is operationalized as the extent to which successive spectral slices differ (or cannot be predicted) from preceding spectral slices. Most simply, CSE is quantified as Euclidean distances between equivalent-rectangular-bandwidth-scaled spectra of fixed-duration (16 ms) sentence slices that were processed by auditory filters" (p. 12388). They report that in their experiment "[t]his pattern of CSE decreasing from low vowels, to high vowels, to laterals/ glides and nasals, to fricatives, to affricates, and finally stops closely parallels the sonority hierarchy" (p. 12389). They conclude that although in their measurements "there are no assumptions that the signal is created by a vocal tract" and "[w]ithout introduction of constructs such as distinctive features, CSE reveals distinctions between classes of speech sounds that mirror those in the sonority hierarchy. Although the present results are agnostic to linguistic constructs and theory, CSE could be construed as providing a case in which some linguistic distinctions are natural consequences of operating characteristics of the mammalian auditory system" (p. 12390).[37]

This is not to say that there is no slack, of course. It is not the case that all languages build the same sonority hierarchy based on the phonetic signal: in some languages, nasals count as (pattern with) sonorants, while in others they go along with stops. Only liquids may be second members of branching onsets in some languages, while in others nasals or glides also qualify for this position. In some languages, only vowels may be sonority peaks, while in others sonorants (or a subset thereof, typically liquids) may also occur in this position and act like vowels (syllabic consonants). The slack that is observed cross-linguistically for the specific location of a given phonetic item on the sonority hierarchy is called 'relative sonority'. This variation has produced a massive body of literature (Clements 1990; Parker 2011, 2017).

Note that the same workings allowing for some slack in a basic invariant frame are known from other cases, where a real-world continuum is converted into discrete cognitive categories. In colour perception, some distinctions like black and white have a physiological basis (magna vs. parvo layer cells of the lateral geniculate nucleus located in the thalamus, Kay et al. 2009: 26) and may be selectively impacted

---

[37]Sign language is argued to have an equivalent of sonority and syllable structure: see for instance Perlmutter (1992) and Sandler (1993). The latter author says that "some kind of movement is necessary for a well-formed sign, just as a vowel, or, in its absence, some other sonorous element, is necessary for a well-formed syllable in spoken language. It has also been noted that movement is perceptually salient, just as vowels are perceptually salient in spoken language" (Sandler 1993: 253). Son in the vocal (and movement in the signed) modality could thus be instantiations of a module whose modality-unspecific domain is perceptual salience. Berent et al. (2013) adduce experimental evidence supporting this amodal view on sonority. Sandler and Lillo-Martin (2006: 235–245) offer an overview of the question. The issue also relates to the notion of headhood, which in Dependency Phonology is reflected by perceptual salience (Anderson and Ewen 1987: 126ff; Anderson 1992: 40ff, 52, 2006: 616).

by dysfunction: there are humans with no chromatic vision (achromatipsia, i.e., seeing only black and white). Beyond physiological grounding, the World Color Survey based on colour naming in 110 languages (Kay et al. 2009: 25) has found that there are six universal sensations (four chromatic: red, yellow, green, blue and two achromatic: black and white) related to colour along which humans partition the colour space.

This partition follows some rules that appear to be universal, but allow for slack: languages may choose to follow a number of different partition paths. Thus, if a language has only two colour-terms, they will partition the spectrum into two chunks whose foci are black (or dark) and white (or light). They are never, say, green and yellow, or black and blue, etc. If there are three colour terms, the third will refer to red/yellow, and in cases where there are four words, the fourth will either split red into red/yellow and green/blue, or into red and yellow (green and blue being versions of black), or into red and yellow/green/blue. Languages with more colour terms also obey this kind of pattern, but the parametrical possibilities increase (Kay and McDaniel 1978, Kay et al. 2009).

Coming back to sonority, this suggests that the conversion of the phonetic signal into Son primes is partly shared by all humans, but allows for some slack that is conventionalized by each language. The work by Iris Berent supports the universal and innate character of sonority. Showing that sonority sequencing is ubiquitous in productive phonological processes, that it is supported by typological data and constrains the behaviour of speakers in psycholinguistic experiments, Berent (2013: 165ff) concludes that sonority sequencing is a grammatical universal since it cannot be derived from extra-grammatical factors (such as phonetics) (see also Berent et al. 2007). She further shows that sonority sequencing does not merely extend to lexical items that speakers have never come across: it is also active in structures that are unattested in the speaker's language, such as branching onset preferences produced by Korean speakers, whose language lacks branching onsets (Berent et al. 2008). Finally, as was mentioned in footnote 37, Berent et al. (2013) adduce experimental evidence for sonority being in fact amodal, that is, a single system shared by the vocal and signed modalities whose expression is loudness etc. on the former, movement on the latter side.

5.2.5.   *Son primes do, Place and Lar primes do not come with a phonetic category*

All representatives of SFP (see footnote 6) except those located at Concordia / Montreal hold that primes are emergent, that is, absent at birth, and that their association with phonetic categories is learned during first language acquisition (2c). The work done at Concordia rejects (2c), arguing that both primes and their association to phonetic categories are genetically coded and present at birth: there is a universal set of substance-free primes (α, β, γ) from which L1 learners choose when building their system, and which are associated to specific phonetic categories (Hale et al. 2007: 647ff, Volenec and Reiss 2018, Reiss and Volenec 2022).[38]

---

[38] The contrast between the two versions of SFP is further discussed in Scheer (2019a: 111–113) and Samuels et al. 2022.

The former option is shown under (10a), the latter under (10b).

(10)   substance-free melodic primes

	phonology	mapping	phonetics
	α	⟷	[x]
	β	⟷	[y]
	γ	⟷	[z]
	melodic primes present at birth?	associations present at birth?	
a. SFP except Concordia	no	no	
b. Concordia	yes	yes	

This distinction corresponds to the difference between the two types of primes discussed in the previous sections: while Place and Lar primes instantiate (10a), Son primes follow (10b).
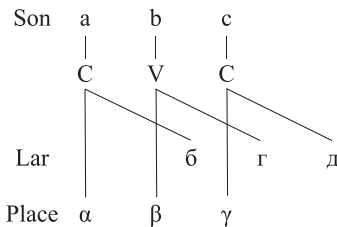
5.3.   Computation upon lexicalization

The existence of three distinct computational systems raises the question of how primes that belong to either can concur in the pronunciation of a single segment: what is the glue that holds the three sets of primes together? The answer can only be the skeleton: a segment is an item that is made of three distinct vocabulary sets which belong to the same timing unit. As is discussed in section 5.4.4 below, these timing units must be specified for consonant- and vowelhood: they therefore appear as "C" and "V" under (11) below.

The linearized string of segments thus defined is the output of the process that converts real world items (the acoustic signal) into cognitive categories (segments) upon lexicalization. The first version of this situation under (5) in section 3 is completed under (11a) with the presence of three distinct vocabulary sets (indicated by three types of symbols; see section 5.2.1): Son, Place and Lar.

(11)   lexicalization
          a. linearized string of segments   b. stored item
             (product of lexicalization conversion)   (product of lexicalization phonology)
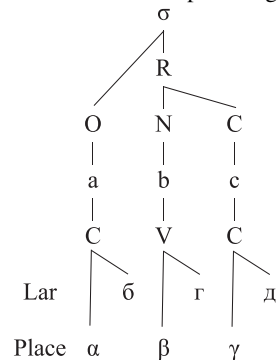
In the structure under (11a), the result of lexicalization conversion are bundles of primes that are assigned to C- and V-slots; each slot together with its primes represents a segment. But the primes are unordered for the time being, since they have not yet undergone any computation in their respective module. This computation (lexicalization phonology) takes (11a) as an input and produces (11b), the object that will be stored in the lexicon (long term memory). The computation in each system is based on relevant primes and projects structure: Son primes a, b, c project structSon (i.e., syllable structure, feet etc.), Lar primes б, г, д project structLar for each segment, and Place primes α, β, γ project structPlace for each segment. What exactly the structure projected in each module looks like varies across theories, but as far as I can see all theories do warrant some organization of primes into structure (rather than having an unordered sets of primes). For example, Feature Geometry organizes primes into a feature geometric structure according to natural classes, while Dependency and Government Phonology assign head or dependent status to primes.[39]

An interesting property of structSon is that it spans segments (a rhyme accommodates two or three segments, a branching onset is made of two segments, government and licensing relate different segments). By contrast, structPlace and structLar are bound by the segment.[40]

5.4.   Computation upon production

This section describes the workings of computation in production, that is, when a string made of lexical items is pieced together and submitted to production phonology.

5.4.1.   Computational domains

In production, a computation is carried out in each module, based on the content of the lexical items (11b) that are retrieved from long-term memory, that is, the linearized string of morphemes that co-occur in a given computational domain (cycle, phase).

Upon production, computation thus takes structure (created upon lexicalization for each morpheme) as an input and returns (modified) structure. The difference between computation creating structure based on primes (lexicalization) and computation modifying existing structure (production) is reminiscent of the distinction that is made in syntax between internal Merge (creation of hierarchy based on primes) and external Merge (movement, i.e., the modification of existing structure).[41]

---

[39]In SFP, "natural" classes are not natural (i.e., phonetically defined) but rather group together primes that share phonological properties in a given language (phonologically active classes, in Mielke's 2008 terms).

[40]A segment as a whole may, of course, be associated to more than one timing unit (long vowels, geminates; see section 5.5.2). But a given segment has only one structPlace and one structLar (no matter how many timing units it belongs to).

[41]Although it is true, of course, that in syntax both types of Merge occur upon production.

In all cases, the domain of phonological computation is defined by the spell-out of morpho-syntactic structure which delineates specific chunks of the linear string that are called cycles or phases. It is obvious and consensual that these chunks are not defined in the phonology.[42]

5.4.2.   StructSon

Computation modifying structSon includes all regular syllable-based processes, stress, as well as the communication with morpho-syntax: structSon (but not structPlace or structLar) and morpho-syntax do communicate, in both directions. Before proper phonological computation can run, though, the string of morphemes over which it operates must be pieced together and linearized.

Computation of structSon caused by linearization is shown under (12a). Morphemes need to be linearized, and this process may be conditioned by phonological information pertaining to Son (infixation (6a)). But items other than morphemes also need to be linearized: there is reason to believe that the exponent of stress may be syllabic space (depending on theoretical inclinations, an x-slot, an empty CV unit, a mora, etc.: Chierchia 1986, Ségéral and Scheer 2008b, Bucci 2013). This extra space thus needs to be inserted to the left or right of the tonic vowel once the location of stress (ictus) is determined. This implies a modification of the linear order present in the lexicon.

The same goes for approaches where the carrier of morpho-syntactic information in phonology is syllabic space: in Strict CV the beginning of the word (or phase) incarnates as a CV unit (the initial CV; see Lowenstamm 1999; Scheer 2012a, 2014a). Note that there are no cases on record where the exponent of morpho-syntactic information is reported to be a Place or a Lar item inserted into structPlace or structLar (such as a labial or a voicing prime at the beginning of the word; see Scheer 2011: Section 663).

In the reverse direction, bearing of structSon on morpho-syntactic computation is documented for the items mentioned under (12b). Again, there are no cases on record where Place, Lar, structPlace or structLar would influence morpho-syntax. That is, morpho-syntax communicates with structSon in both directions, but is completely incommunicado with Place or structPlace, Lar or structLar (Scheer 2012a: Section126, see footnote 27).

Finally, cases of regular phonological computation involving structSon are shown under (12c).

---

[42]The Prosodic Hierarchy (Selkirk 1981, Nespor and Vogel 1986) is a representational way of defining phonologically relevant chunks. The traditional assumption is that phonologically relevant chunks are derivationally defined below the word level (cycles), but have a representational definition at and above the word level (prosodic constituency). This peaceful coexistence (Scheer 2011: Sections 435–440) is called into question by phase theory (D'Alessandro and Scheer 2015), but the debate is orthogonal to the present discussion.

(12) Computation of structSon upon production

    a. linearization

       the linear sequence that undergoes phonological computation is pieced together from various sources:

       1. the lexical entries of morphemes that occur in the computational domain: affix placement, including infixation (definition of the landing site, see (6a))

       2. insertion of the exponent of stress which may be syllabic space (an x-slot, an empty CV unit, a mora, etc.).

       3. exponents of morpho-syntactic information such as the initial CV representing the beginning of the word (or phase).

    b. structSon information that is used by morpho-syntactic computation

       1. phonologically conditioned allomorphy; see (6b)

       2. intonation and stress
Szendrői (2003, 2004); Hargus (1993)

       3. tone
Rose and Jenks (2011)

       4. size of lexical items, for example, minimal word constraints: number of syllables or moras
Inkelas and Zec (1990: 372ff), Hargus (1993)

       5. rhythm
 Guasti and Nespor (1999)

    c. phonological computation

       (that may or may not use morpho-syntactic information)

       1. stress placement (in languages where stress is not lexical)

       2. calculation of syllable structure
across the syllabic material that is pieced together from different sources (12a)

       3. calculation of positional strength
the (relative) strength of positions depends on syllable structure.

       4. lenition and fortition
modification of primes due to structSon: for example, lenition in coda position such as ł → w modifies relevant Son primes, turning a lateral into a glide in the specific structSon configuration "coda". This pattern is discussed in section 4.4.

5.4.3.   structPlace and structLar

There is also a computation concerning Place primes and structPlace, such as palatalization, vowel harmony etc.: the prime associated to palatality is absent in velar but present in palatal segments, and a palatal source may add palatality to velars (front vowel harmony in V-to-V interaction, palatalization in V-to-C interaction). There is no need for going into more detail regarding Place and Lar: relevant phenomena are well known and not any different in the present environment from what they are in regular autosegmental structure: spreading, linking and delinking of primes.

5.4.4.   Timing units are specified for consonant- and vowelhood

As is obvious in general and from the preceding, Place computation is informed of whether the items computed belong to a consonant or a vowel: vowel harmony is only among vowels, and in regular palatalization the trigger is a vowel and the

target, a consonant. Information about vowel- and consonanthood is absent from the Place system, though. It must therefore be present in the only other structure that Place has access to: the timing units that represent linearity.

That Place has access to consonant- and vowelhood is also shown by a fundamental insight of autosegmentalism: the same primes may produce different segments depending on whether they are associated to consonantal or vocalic slots. High vowels and corresponding glides i-j, u-w, y-ɥ are the same segmental items and distinct only through their affiliation to a consonantal or a vocalic slot (Kaye and Lowenstamm 1984). For instance, when an i spreads its segmental content to a vacant consonantal position, it is interpreted as j: French *li-er* "to link" is pronounced [li-j-e], but the j is absent from both the root (*il lie* [li] "he links") and the infinitive suffix (*parl-er* [paʁl-e] "to talk"). The same goes for *lou-er* "to rent" [lu-w-e] and *tu-er* "to kill" [ty-ɥ-e]: in all cases the glide is a copy of the preceding vowel into a C-position. This is shown under (13) where the Element I represents the Place specification of i/j (in substance-laden guise).

(13) high vowels and corresponding glides are segmentally identical
       French *li-er* [li-j-e] "to link"
       a. lexical ingredients   b. spreading
       C V - C V   C V C V
       | | | | ↗ |
       l I   e   l I   e   [lije]

The idea that timing units are specified for consonant- and vowelhood goes back to Clements and Keyser (1983).

5.5. Spell-out

This section discusses the question of which unit exactly is spelled out. It is concluded that segment integrity allows for only one solution: the root node (in the sense of Feature Geometry).

5.5.1. *Multiple-module spell-out*

The existence of several computational systems (Son, Place and Lar) that concur in the pronunciation of a single segment raises a non-trivial challenge for the spell-out mechanism at the phonology–phonetics interface. Note that the situation is quite different from what we know from the morpho-syntax–phonology interface where the input to spell-out comes from one single system (morpho-syntax) and hence conveys one single vocabulary.

At the phonology–phonetics interface, Son, Place and Lar do not see each other and cannot communicate, but are segment-bound, that is, they must co-define the same segment. Therefore, segment integrity can only exist if the items of the three computational systems are related through a fourth player that defines the segment: the skeleton. Timing units that embody the linear structure of the string do precisely

the job that is needed, that is, the definition of what does and does not belong to a segment. The difference between the upper spell-out that receives information from one single module and the lower spell-out which needs to combine inputs from three different modules is thus reflected by the presence of linearity in the latter, against its absence in the former. It is only the existence of linearity and hence the skeleton that affords the existence of a multiple-module spell-out.

It follows that the lower interface does not spell out individual primes or the set of primes defined by each one of the three computational systems, but entire segments as defined by timing units. How exactly timing units relate to segments is discussed in the following section.
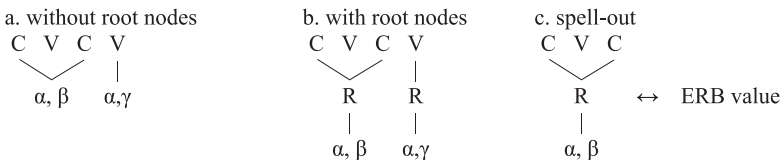
5.5.2. *Segment-defining root nodes are spelt out*

Regular autosegmental representations allow for the association of a set of primes to multiple timing units as under (14a) where a geminate is shown. It is obvious that the primes must "know" whether they are associated to one or two timing units: this makes the difference between a short and a long item.

Let us consider the spell-out algorithm, that is, the way units are defined that are submitted to phonetic conversion. Spell-out cannot proceed timing unit by timing unit (Cs and Vs under (14)) since this would create two identical phonetic items under (14a): the association of α, β to a C would be spelt out twice, for the two Cs of the geminate. What is realized, though, is not two times the consonant in question, but the consonant with a longer duration (or with other phonetic correlates; see (19) below).[43]

Spell-out cannot proceed prime by prime either, since in this case α of the geminate under (14a) will be converted into a phonetic item, followed by the conversion of β, effecting, on the phonetic side, a sequence of items. In this case the phonetics would not "know" that the representatives of α and β need to be pronounced simultaneously, rather than in sequence.

(14) regular autosegmental representations

a. without root nodes
```
C   V   C   V
 \_/      |
 α, β   α,γ
```

b. with root nodes
```
C   V   C   V
 \_/      |
  R       R
  |       |
 α, β   α,γ
```

c. spell-out
```
C   V   C
 \_/
  R      ↔   ERB value
  |
 α, β
```

It thus appears that the representation under (14a) is not suited to identifying the geminate segment at hand, that is, "α and β associated to two consonantal timing units": there is no single item that encompasses this set. In traditional Feature

---

[43]This is the case of true geminates (or long vowels). There are also fake geminates, that is, the consecution of two identical consonants that are each associated to a timing unit (Hayes 1986).

Geometry (Clements and Hume 1995), the segmental unit is defined by the root node. This option is shown under (14b) where "R" is the root node.[44] If the segment is defined by this item, spell-out may proceed root node by root node: for every root node present on the phonological side, a spell-out instruction is searched that contains all items affiliated. That is, the geminate under (14b) will be able to be spelt out by the spell-out instruction under (14c) (phonetic categories appear as ERB values; see section 5.6.2): spell-out will transmit the phonetic correlates of the primes contained in the root node (α, β), as well as the phonetic correlate (duration) corresponding to the fact that the root node is attached to two timing units (length).

This means that languages possess exactly as many spell-out instructions as there are phonemes (segments), that is, objects that are phonologically distinct. In comparison to the upper interface of morpho-syntax and phonology, the number of lexical entries whose content is inserted into the receiving module is thus very small: thousands of morphemes against 15 or 20, in very large inventories maybe 40 phonemes (segments).

5.5.3.   *Workings*

When phonological computation (in all three modules) is completed, the result is spelled out root node by root node (i.e., segment by segment): root nodes are associated to Son, Place and Lar primes, which in turn are associated to their respective structure (see (11)). Spell-out considers only timing units and primes (Son, Place, Lar) associated to a given root node. The structure projected by primes (structSon, structPlace, structLar) is not taken into account. The restriction of spell-out to timing units and primes is motivated in section 6 below.

Spell-out matches the output of phonological computation with the phonological side of spell-out instructions, a lexicon where (language-specific and hence acquired) correspondences between phonological and phonetic categories are defined. For example, the spell-out lexicon entry under (15a) associates the Place prime α and the Son prime b, both belonging to a root node "R" and linked to a V slot, with the phonetic value "high front unrounded vowel" (or rather, the corresponding ERB value; see section 5.6.2). This assures the spell-out of the leftmost V of French *li-er* [lije] "to link" under (13b). By contrast, the second C under (13b) will match the spell-out entry (15b) where the same primes α (Place) and b (Son) are associated to a C position, a configuration whose phonetic correspondence is a high front unrounded glide.

---

[44]Note that timing units (x-slots, Cs and Vs under (14), moras etc.) are not the same thing as root nodes. The latter have no timing properties but rather define segments. That timing units and segments may not coincide is shown by the present discussion: the segment that makes a geminate cannot be characterized by any given timing unit. Further arguments for the root node come from floating segments, as shown by Cavirani (To appear).

(15) spell-out lexicon
    a. V-R-α,b ↔ high front unrounded vowel
    b. C-R-α-b ↔ high front unrounded glide

5.5.4. Spell-out mismatches

Spell-out is a list-based mapping of one vocabulary set onto another. Therefore, the relationship between the two items matched is necessarily arbitrary (2c). The overwhelming majority of mappings appears to be "natural", that is, phonetically reasonable, though: an item that shows the phonological behaviour of a labial is also realized as a phonetic labial. The overwhelming naturalness of mappings is due to the fact that phonological categories come into being when a phonetic item or event is phonologized. At birth, phonology is thus phonetically transparent and natural, with a faithful mapping. It takes some accident in the further life of phonological processes for them to become crazy, or for phonological items to develop a non-faithful mapping (Bach and Harms 1972, Scheer 2014b: 268ff, Chabot 2019). That is, processes are not born crazy and mappings are not born unfaithful – they may become crazy and unfaithful through aging. This is one aspect of the life-cycle of phonological processes (Baudoin de Courtenay 1895, Vennemann 1972, Bermúdez-Otero 2007, 2015).

Nonetheless, nothing in the system enforces or favours faithful mappings: "naturalness" and faithfulness are imposed by workings that lie outside both phonology and spell-out (third factor in the sense of Chomsky 2005; see Chabot 2021). Spell-out is happy to accommodate any type of mapping, faithful or arbitrary. The ability of the system to lexicalize arbitrary mappings should thus have some empirical echo: at least some phonology–phonetics mismatches should exist. The cases mentioned under (16)–(19) show that this is indeed the case for all components discussed: Son under (16), Place under (17), Lar under (18) and timing units under (19). Note that the phenomena mentioned for the sake of illustration are only a small subset of the empirical record: phonology–phonetics mismatches are pervasive cross-linguistically (Hamann 2011, 2014).

(16) spell-out mismatches: Son
    a. the phonological sonorant r appears as ʃ/ʒ in Polish, as ʁ/χ in French (Chabot 2019).
    b. the phonological sonorant w has an obstruent realization v in a number of Slavic languages (Cyran and Nilsson 1998).

(17) spell-out mismatches: Place
    uu-fronting in South-East British English
    boot is pronounced [biit] (Henton 1983, Harrington et al. 2008), but [ii] continues to trigger w-insertion in *do it* [dii w it], thereby attesting its phonological identity /uu/.

(18) spell-out mismatches: Lar
       preaspiration
       cross-linguistically, preaspiration realizes what were formerly phonetic geminates
       (Clayton 2010). In Icelandic for example (Hansson 2001, Helgason 2002: 216ff),
       voiceless geminates are realized as preaspirated but without extra duration (i.e., like non-
       geminates). Thus Old Norse p, t, k are represented as such, Old Norse voiced geminates
       bb, dd, gg appear as voiceless geminates pp, tt, kk, but Old Norse voiceless geminates are
       preaspirated ʰp, ʰt, ʰk. Beyond the diachronic patterning, the geminacy of ʰp, ʰt, ʰk is
       evidenced by their synchronic behaviour (long vowels cannot appear to their left).

(19) spell-out mismatches: timing units (length)
       In the case of timing units, the length of phonologically long vowels and phonological
       geminates may be marked in the phonetic signal by duration (no mismatch), but also by
       other means (mismatch). In case their long identity is flagged by some other property,
       their own realization or the context allows the speaker (and learner) to recover underlying
       length.[45],[46]
       a. In English and German, agma [ŋ] is a singleton on the surface, but represents the
            partial geminate /ŋg/: diagnostics are the fact that it occurs only after short vowels
            and is absent word-initially (Dressler 1981, Gussmann 1998).
       b. Hammond (1997) argues that in English, intervocalic (singleton) consonants are
            phonological geminates when preceded by a short/lax vowel (e.g., the *t* in *city*). This
            is because short/lax vowels occur only in closed syllables: the i in *city* must thus be
            checked.[47]
       c. Virtual long vowels
            rather than by duration, the phonological contrast between long and short vowels
            was found to be phonetically expressed by
            1. reduced (schwa, representing short items) vs. unreduced (non-central vowels,
                 representing long items) vowels in Semitic (Lowenstamm 1991, 2011), Ge'ez
                 (Old Ethiopian) (Ségéral 1996) and Kabyle Berber (Bendjaballah 2001, Ben Si
                 Saïd 2011).
            2. ATRness in French (+ATR representing long, -ATR short mid vowels; see
                 Rizzolo 2002).

---

[45]Especially in case the only cue to geminacy is found in the surrounding context, non-
executed duration may be interpreted as a means of realizing economy: if geminacy is
already marked, when for example on the preceding vowel, there is no need to realize it a
second time on the consonant.

[46]An interesting question regarding timing mismatches is why phonological length may be
pronounced without duration (/CC/ ↔ [C], /VV/ ↔ [V]), while the reverse pattern where a
phonologically short item is phonetically realized with extra duration (/C/ ↔ [CC], /V/ ↔
[VV]) appears to be absent from the record.

[47]Geminates and so-called ambisyllabic consonants (Kahn 1980 and ensuing literature)
both belong to two syllables and are therefore heterosyllabic. The difference is that the
former are associated to two, the latter only to one, timing unit. Ambisyllabicity is the represen-
tation developed in times when phonology–phonetics mismatches could not be envisioned, that
is, where a phonetically simplex consonant could not possibly belong to two timing units.

  d. Virtual geminates
      rather than by duration, phonological geminates were found to be expressed by
      1. the length of the preceding vowel in German (Caratini 2009), the Cologne dialect of German (Ségéral and Scheer 2001b) and English (Hammond 1997, see (19b)).
      2. the inhibition of a preceding vowel-zero alternation in Somali (Barillot and Ségéral 2005).
      3. preaspiration, see (18).

5.6. Phonetics

In this section it is argued that phonology is spelt out to a language-specific phonetic system (LSP) that, following the BiPhon model, is acoustic in kind and works on a vocabulary made of ERB values. LSP is followed by articulatory universal phonetics.

5.6.1. Language-specific phonetics

There is reason to believe that phonetics falls into two distinct computations: one that is language-specific, acquired during L1 acquisition and cognitive in kind (Language-Specific Phonetics, LSP), the other being universal and located outside of the cognitive system (Universal Phonetics). While LSP is acoustic in kind and defines language-specific gradient properties that identify a specific dialect or pronunciation (e.g., *t* being dental, alveolar, post-alveolar etc.), Universal Phonetics is shared by all languages: it is articulatory in kind and based on coarticulation, aerodynamics and other physical / physiological properties.

Unlike phonology which manipulates discrete objects, both phonetic systems are gradient in kind. LSP thus identifies as a cognitive system which is part of the grammar that speakers acquire, and like other grammatical systems (phonology, syntax) affords well-formedness: phonetic items may be well- or ill-formed (this is expressed by articulatory constraints in BiPhon; see Boersma and Hamann 2008: 227ff). Kingston (2019: 389) says that "speakers and listeners have learned a *phonetic* grammar along with their phonological grammar in the course of acquiring competence in their language" (emphasis in original).

Evidence for LSP comes from the fact that "speakers' phonetic behaviour cannot be entirely predicted by the physical and physiological constraints on articulations and on the transduction of those articulations into the acoustic properties of the speech signal" (Kingston 2019: 389). A classical case is vowel duration, which is variable according to the voicing of the following consonant: voiced articulations (both sonorants and voiced obstruents) provoke longer durations to their left than voiceless obstruents. This appears to be (near) universal (Chen 1970), but the ratios of long and short vowels are quite different across languages: Cohn (1998: 26) reports that while vowels before voiceless consonants are 79% of the length of vowels before voiced consonants in English, in Polish the ratio is 99%, with other languages coming in between these values. This suggests that some phonetic properties are under the control of the speaker and depend on the language they have acquired.

Relevant literature describing the properties and workings of LSP includes Keating (1985), Pierrehumbert and Beckman (1988), Cohn (1998), Cho and Ladefoged (1999), Boersma et al. (2003); Cho (2011: 343-346), and Kingston (2019) provide overviews.

5.6.2.   Phonetic categories

If a spell-out hands over the output of phonology to another computational system that is cognitive (and grammatical) in kind, LSP, the question arises: what does the domain-specific vocabulary look like in this system? In the BiPhon model, the auditory continuum is expressed as an auditory spectral mean along the ERB scale (Equivalent Rectangular Bandwidth) (Boersma and Hamann 2008: 229).

That is, the spell-out from phonology to LSP associates a phonological and a phonetic category, for example s ↔ 20,1 ERB. This reads "*s* is realized with a spectral mean of 20,1 ERB" and in a substance-free environment the consonant is replaced by alphas and betas. The goal of LSP (Auditory Form in BiPhon) in production is to determine an acoustic target in form of an ERB value that is transmitted to Universal Phonetics (Articulatory Form in BiPhon). Since LSP is gradual, but a simple spell-out instruction such as s ↔ 20,1 ERB produces a discrete ERB value, the computation carried out in the LSP module creates relevant variation: given an invariable and discrete phonological *s*, LSP computation produces ERB values with slight variation around the input from the spell-out instruction. This is responsible for some of the phonetically variable realizations of *s*. Variation/gradience is thus introduced by both the variable acoustic target that LSP produces and transmits to Universal Phonetics and the undershoot/overshoot that this target is subject to in the physical implementation carried out by Universal Phonetics.

Note that the red line between the discrete and the gradual is thus crossed in LSP: the input to LSP is a discrete ERB value contained in each spell-out instruction (s ↔ 20,1 ERB etc.). The modular computation operates over this ERB value and produces a gradient output. In the BiPhon model, the generation of the gradient output is achieved by adding noise to the computation either in the guise of multiple, slightly varying ERB values for a given spell-out that undergo an OT computation (constraints such as *s ↔ 20,1 ERB, *s ↔ 20,2 ERB, *s ↔ 20,3 ERB etc., Boersma and Escudero 2004, Boersma and Hamann 2008: 233ff) or by having the computation done by an artificial neural network (Seinhorst et al. 2019).

5.6.3.   Visibility of morpho-syntactic divisions

Since LSP is a grammar-internal computational system, like all other modules it computes specific stretches of the linear string that are defined in terms of morpho-syntactic divisions, (i.e., cycles or phases; see section 5.4.1).

The visibility of morpho-syntactic divisions for phonetic processes is a disputed issue: in a feed-forward implementation of modularity as argued for by Bermúdez-Otero and Trousdale (2012), modules can only take into account information of the immediately preceding module. Since phonology intervenes between phonetics and morpho-syntax, this perspective predicts that morpho-syntactic information

will never be taken into account by phonetic processes. In a diachronic perspective, on this count, it is only when phonetic processes have been phonologized that a morpho-syntactic conditioning may kick in.

In a regular modular architecture (see section 2.4.2), there is no reason to restrict the availability of morpho-syntactic divisions, though. Like all other modules, LSP necessarily applies to a given computational domain. It is unclear how a cognitive computational system could work at all in absence of a specified stretch of the linear string over which it operates: there is no computation in absence of a computational domain. Hence the chunks defined in the morpho-syntax are handed down to all subsequent computational systems until the verge of the cognitive system is reached: universal phonetics is not cognitive in kind, and hence unbound by computational domains.[48]

Starting with Lehiste (1960), there is a substantial literature documenting cases where morpho-syntactic divisions are taken into account by phonetic processes. A well-studied item is *l*-darkening in English (Giles and Moll 1975, Lee-Kim et al. 2013, Strycharczuk and Scobbie 2016, Mackenzie et al. 2018), but there is debate as to whether the gradient darkening overlaps with a categorical phonological process (Bermúdez-Otero and Trousdale 2012, Turton 2017). If this is the case, Bermúdez-Otero and Trousdale (2012), who reject the visibility of morpho-syntactic information in phonetics, can argue that *l*-darkening is influenced by morpho-syntactic divisions not in the phonetics, but in its phonological incarnation.[49]

In order to get around this caveat, Strycharczuk and Scobbie (2016) have tested whether phonetic processes may be sensitive to morpho-syntactic divisions. They study fronting of the *goose* vowel [uu] (see (17) in section 5.5.4), an ongoing sound change in Southern British English that is reported to be inhibited by a following coda *l* as in *fool*. The authors have experimentally contrasted *fool* with intervocalic *l* that is (*fool-ing*) or is not (*hula*) morpheme-final. Their results show that fronting is more inhibited in the former than in the latter case, thus documenting the impact of the morphological boundary. They carefully argue that the process at

---

[48]The literature on production planning (Wagner 2012, Tanner et al. 2017, Kilbourn-Ceron and Sonderegger 2018, Tamminga 2018) works with a production planning window (or production scope) that defines the stretch of the linear string for which production is prepared in one go. This window is variable (across speakers, individual speech acts etc.) and defined by a number of factors that include morpho-syntactic information. That is, a production planning window may not be identical to a phase (or a cycle), but will in part have been defined by it.

[49]From the point of view of perception, the question is whether phonetic information is used in order to identify morpho-syntactic divisions. There is certainly good reason to assume that any information available may be used in order to parse the phonetic signal. In functional, listener-oriented work such as Boersma (2007), speakers in fact *create* relevant phonetic and phonological information "on purpose" in order to enhance perception. But this does not speak to the issue at hand, since Bermúdez-Otero and Trousdale's (2012) feed-forward model is about production and does not object to phonetic information being used for the identification of morpho-syntactic divisions in perception. But even if it did, the question would be the same as in production: maybe the transmission is not direct but through the phonology, which then complies with feed-forward standards.

hand is gradient, rather than categorical, and therefore not a case of phonologically controlled allomorphy. They conclude that "morphological boundaries may affect phenomena that are phonetically continuous and gradient, and not only clear cases of allophony" (Strycharczuk and Scobbie 2016: 90).

6.   IS A GIVEN ALTERNATION COMPUTATIONAL OR INTERPRETATIONAL IN KIND?

Approaches that provide for a spell-out operation in addition to regular phonological computation are confronted with the question whether a given modification of the lexical representation is computational (caused by phonological computation) or interpretational (due to spell-out) in kind.

   Consider the trivial case of *l*-vocalization whereby lexical *l* (or *ł*) appears as *w* in coda position (for example in Brazilian Portuguese, Collischonn and Costa 2003). This could be due to phonological computation (Son primes are rearranged under positional pressure), or to a spell-out instruction "l (coda) ↔ w". On the latter count, the phonological side of the spell-out instruction not only mentions Son primes (the definition of a lateral), but also structSon, that is, the fact for the lateral to belong to a coda position.

   This would mean that lenition and fortition are subjected to the arbitrariness of spell-out (2c). Thus the existence of mismatches (or crazy matches) that are typical for spell-out relations (section 5.5.4) is expected: laterals could appear as θ in coda position (l (coda) ↔ θ), or they could appear as *w* in strong word-initial position (anti-*l* vocalization: l (word-initial) ↔ w). Such mismatches do not appear to exist and, significantly, Chabot's (2021) inventory of crazy rules (section 2.3) features no cases of a syllabic conditioning: i → u / d__ is reported for Southern Pomoan, but there is no case of, say, i → u in open syllables. This suggests that structSon is not used in spell-out instructions.

   Thus there is reason to believe that, at least for Son, only primes are spelled out: structure is absent from spell-out. In our example, then, *l*-vocalization in coda position is effected by regular phonological computation of Son. Like all other segments, the result of this computation, *w*, undergoes spell-out and may be subject to a mismatch, say, w ↔ θ. But this will then concern *all w*'s of the language, not just those that occur in a coda. By contrast, if spell-out instructions were sensitive to syllable structure, lexical laterals could appear as θ only in codas through a spell-out instruction "l (coda) ↔ θ". It was mentioned that this kind of syllable-sensitive crazy rule appears to be absent from the record.

   This ties in with the insight from section 2.4.1: spelling out an item is the process of assigning it a phonetic value. Items of structSon do not have any phonetic correlate, though, and therefore cannot be spelled out. Their absence from spell-out instructions is also more generally consistent with the fact that processes at and above the skeleton are not arbitrary in the empirical record (section 2.3): their computation is not arbitrary because they are based on meaningful vocabulary, (2b). In sum, Son primes may experience arbitrary distortion (through spell-out), but structSon may not: it always reaches (language-specific) phonetics faithfully.

Note, however, that illusions of arbitrary Son computation may be produced by the system. Suppose an underlying *b* is turned into *β* by Son computation and this *β* is then spelled out as, say, *θ*. This creates the impression of an arbitrary computation b > θ, but in fact involves regular computation followed by a spell-out mismatch.

7.   CONCLUSION

The architecture discussed prompts some more general questions that are addressed below.

Phonological computation occurs twice, upon lexicalization and upon production. Syllabification, for example, runs on both occasions, the difference being the portion of the linear string that is concerned: first within the to-be-lexicalized morpheme (lexicalization), then over the plurimorphemic domain defined by morpho-syntax, that is, the cycle/phase (production).

This may be construed as an instance of the so-called Duplication Problem, a term coined by OT that describes a situation where the same labour (here syllabification) is done twice, in the lexicon and upon production. In OT, Morpheme Structure Constraints (MSC, mentioned in section 3) have motivated the principle of Richness of the Base (Prince and Smolensky 2004: 191, McCarthy 1998): as McCarthy (2003: 29) puts it, "OT solves the Duplication Problem by denying the existence of morpheme structure constraints or other language-particular restrictions on underlying forms. OT derives all linguistically significant patterns from constraints on outputs interacting with faithfulness constraints ('Richness of the Base' in Prince and Smolensky 2004)". Vaux (2005: 5) believes that "the duplication argument […] is the heart of the attack on MSCs and in general perhaps the most invoked OT argument against DP [Derivational Phonology]". Morpheme Structure Constraints are restrictions on lexical representations, such as the prohibition of morpheme-initial sonorant-obstruent clusters in English, which would also be an impossible output of phonological computation upon production (the restriction is stated twice, in the lexicon and in computation). Vaux (2005), Rasin and Katzir (2015), Rasin (2018: 93-152), argue that there is nothing wrong with this kind of redundancy because MSCs are conceptually and empirically necessary. The architecture exposed in this article is in line with this position.

Another question is why Son is related to morpho-syntax (in both directions, section 5.4.2), while Place and Lar are not: they are incommunicado with morpho-syntax in both directions (sections 4.1.3 and 5.4.2, footnote 27). There is no answer to this question other than the observation that the wiring among modules is more generally unpredictable. The McGurk effect (McGurk and MacDonald 1976) for example documents that vision influences Place: in so-called McGurk fusion, subjects who are presented with synchronized visual *g* (the video recording of somebody pronouncing [g]) and audio *b* (the audio recording of somebody pronouncing [b]), perceive *d*. Hence g_{video} and b_{auido} have combined into *d*, even though *d* is absent from the sensory input to the subject. The impact of vision is thus on Place. There is no McGurk effect reported on Son, that is, a situation

where, say, visual *p* (an obstruent) combined with audio [a] produces the perception of *m*, that is, a sonorant that occurs half way on the sonority hierarchy. In the same way, synesthesia relates colours with objects or concepts in an unpredictable way: synesthetic subjects (about 4% of the population) associate this or that colour to this or that number, letter, person, concept, sound, etc. (Sagiv and Ward 2006).[50]

Finally, the question arises as to why Son is phonologically meaningful but Place and Lar are not. One consequence is that there must be a cross-linguistically uniform way for humans to extract Son-relevant properties from the phonetic signal, while the association of Place and Lar to phonetic categories is language-specific and emergent (section 5.2). A possible answer is that the correlate of Son, perceptual salience (loudness, energy and intensity, section 5.2.3), is used for other purposes by humans, while Place and Lar appear to do labour only in phonology. That is, salience is used to evaluate emotions, danger, relevance of noise etc., while Place and Lar distinctions do not play any role for humans outside of phonology: what would a distinction between, say, labial and velar, or voiced and voiceless, be used for outside of speech?

In an evolutionary/biolinguistic perspective, this means that humans have developed a stable means of converting salience into cognitive categories, and Son is an exaptation thereof that occurred when language emerged in the species. In evolution, an exaptation is an opportunistic adaptation of an existing device for a function that it was not designed for. That is, Son primes and their association to phonetic categories predate language in the species and are likely present in animal cognition, while this is not the case for Place and Lar. These are rather a creation *ex nihilo* for the specific needs created by language, that is, the expression of contrast with the anatomic devices offered by speech organs. That is, new cognitive categories representing Place and Lar needed to be created by each language individually and *ex nihilo*: this may be the reason why they are arbitrary (phonologically meaningless), and why their association to phonetic categories is not hard wired but rather language-specific.

REFERENCES

Adger, David, and Peter Svenonius. 2012. Features in minimalist syntax. In *The Oxford handbook of minimalist syntax*, ed. Cedric Boeckx, 27–51. Oxford: Oxford University Press.
Alber, Birgit. 2001. Regional variation and edges: Glottal stop epenthesis and dissimilation in standard and Southern varieties of German. *Zeitschrift für Sprachwissenschaft* 20: 3–41.
Anderson, John. 1992. *Linguistic representation. Structural analogy and stratification*. Berlin: Mouton de Gruyter.
Anderson, John. 2006. Structural analogy and Universal Grammar. *Lingua* 116: 601–633.
Anderson, John. 2011. *The substance of language. Vol.1 The domain of syntax. Vol.2 Morphology, paradigms, and periphrases. Vol.3 Phonology syntax analogies*. Oxford: Oxford University Press.

---

[50]The mutual incommunicadoness between Place, Lar and Son that motivates their status as independent modules is another case in point: there could be a wire between either of them in the form of a modular interface, but the observation is that there isn't.

Anderson, John, and Colin Ewen. 1987. *Principles of dependency phonology*. Cambridge: Cambridge University Press.

Bach, Emmon, and R. T. Harms. 1972. How do languages get crazy rules? In *Linguistic change and generative theory*, ed. Robert Stockwell and Ronald Macaulay, 1–21. Bloomington: Indiana University Press.

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

Backley, Phillip, and Kuniya Nasukawa. 2020. Recursion in melodic-prosodic structure. In *Morpheme-internal recursion in phonology*, ed.Kuniya Nasukawa, 11–35. Berlin: de Gruyter.

Bakst, Sarah, and Jonah Katz. 2014. A phonetic basis for the sonority of [X]. *UC Berkeley Phonology Lab Annual Report* 10: 11–19.

Barillot, Xavier, and Philippe Ségéral. 2005. On phonological processes in the '3rd' conjugation in Somali. *Folia Orientalia* 41: 115–131.

Baudouin de Courtenay, Jan Niecisław. 1895. *Versuch einer Theorie phonetischer alternationen. Ein Capitel aus der Psychophonetik [Attempt at a theory of phonetic alternations. A chapter from psycho-phonetics]*. Straßburg: Trübner.

Beckman, Mary, Jan Edwards, and Janet Fletcher. 1992. Prosodic structure and tempo in a sonority model of articulatory dynamics. In *Gesture, segment, prosody: Papers in laboratory phonology II*, ed. Gerard J. Docherty and D. Robert Ladd, 68–86. Cambridge: Cambridge University Press.

Ben Si Saïd, Samir. 2011. Interaction between structure and melody: The case of Kabyle nouns. In *On words and sounds*, ed. Kamila Dębowska-Kozłowska and Katarzyna Dziubalska-Kołaczyk, 37–48. Newcastle upon Tyne: Cambridge Scholars.

Bendjaballah, Sabrina. 2001. The negative preterite in Kabyle Berber. *Folia Linguistica* 34: 185–223.

Berent, Iris. 2013. *The phonological mind*. Cambridge: Cambridge University Press.

Berent, Iris. 2018. Algebraic phonology. In *The Routledge handbook of phonological theory*, ed. S.J. Hannahs and Anna Bosch, 569–588. Oxford: Routledge.

Berent, Iris, Tracy Lennertz, Jongho Jun, Miguel A. Moreno, and Paul Smolensky. 2008. Language universals in human brains. *Proceedings of the National Academy of Sciences of the United States of America* 105: 5321–5325.

Berent, Iris, Donca Steriade, Tracy Lennertz, and Vered Vaknin. 2007. What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104: 591–630.

Berent, Iris, Amanda Dupuis, and Diane Brentari 2013. Amodal aspects of linguistic design. *Plos One* 8, article e60617.

Bermúdez-Otero, Ricardo. 1999. Constraint interaction in language change: Quantity in English and German. Doctoral dissertation, University of Manchester.

Bermúdez-Otero, Ricardo. 2003. The acquisition of phonological opacity. In *Variation within Optimality Theory: Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, ed. J. Spenader, J. Eriksson and A. Dahl, 25–36. Stockholm: Department of Linguistics, Stockholm University [longer version at ROA #593].

Bermúdez-Otero, Ricardo. 2007. Diachronic phonology. In *The Cambridge handbook of phonology*, ed.Paul de Lacy, 497–518. Cambridge: Cambridge University Press.

Bermúdez-Otero, Ricardo. 2015. Amphichronic explanation and the life cycle of phonological processes. In *The Oxford handbook of historical phonology*, ed.Patrick Honeybone and Joseph C. Salmons, 374–399. Oxford: Oxford University Press.

Bermúdez-Otero, Ricardo, and Graeme Trousdale. 2012. Cycles and continua: On unidirectionality and gradualness in language change. In *The Oxford handbook of the history of*

English, ed.Terttu Nevalainen and Elizabeth Closs Traugott, 691–720. NewYork: Oxford University Press.

Blaho, Sylvia. 2008. The syntax of phonology. A radically substance-free approach. Doctoral dissertation, University of Tromsø.

Blevins, Juliette. 1995. The syllable in phonological theory. In *The handbook of phonological theory*, ed.John Goldsmith, 206–244. Oxford, Cambridge, Mass: Blackwell.

Bloch, Bernard, and George Trager. 1942. *Outline of linguistic analysis*. Baltimore: Linguistic Society of America.

Bloomfield, Leonard. 1939. Menomini morphophonemics. *Travaux du Cercle linguistique de Prague* 8: 105–115.

Bloomfield, Leonard. 1962. *The Moenomini language*. New Haven and London: Yale University Press.

Boersma, Paul. 1998. *Functional phonology. Formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.

Boersma, Paul. 2007. Some listener-oriented accounts of *h-aspiré* in French. *Lingua* 117: 1989–2054.

Boersma, Paul, and Paola Escudero. 2004. Bridging the gap between L2 speech perception research and phonological theory. *Studies in second language acquisition* 26: 551–585.

Boersma, Paul, Paola Escudero, and Rachel Hayes. 2003. Learning abstract phonological from auditory phonetic categories: An integrated model for the acquisition of language-specific sound categories. In *Proceedings of the 15th International Congress of Phonetic Sciences*, ed.Maria-Josep Sole, Daniel Recasens and Joaquín Romero, 1013–1016. Barcelona: Universitat Autónoma de Barcelona.

Boersma, Paul, and Silke Hamann. 2008. The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25: 217–270.

Bucci, Jonathan. 2013. Voyelles longues virtuelles et réduction vocalique en coratin. *Canadian Journal of Linguistics* 58(3): 397–414.

Buckley, Eugene. 2000. On the naturalness of unnatural rules. *UCSB Working Papers in Linguistics* 9.

Buckley, Eugene. 2003. Children's unnatural phonology. *Proceedings of the Berkeley Linguistics Society* 29: 523–534.

Caratini, Emilie. 2009. Vocalic and consonantal quantity in German: Synchronic and diachronic perspectives. Doctoral dissertation, Nice University and Leipzig University.

Carruthers, Peter. 2006. *The architecture of the mind. Massive modularity and the flexibility of thought*. Oxford: Clarendon Press.

de Carvalho, Joaquim Brandão. 2002a. De la syllabation en termes de contours CV. Thèse d'habilitation, École des Hautes Études en Sciences Sociales, Paris.

de Carvalho, Joaquim Brandão. 2002b. Formally-grounded phonology: From constraint-based theories to theory-based constraints. *Studia Linguistica* 56: 227–263.

de Carvalho, Joaquim Brandão. 2008. From positions to transitions: A contour-based account of lenition. In *Lenition and fortition*, ed.Joaquim Brandão de Carvalho, Tobias Scheer and Philippe Ségéral, 415–445. Berlin: de Gruyter.

de Carvalho, Joaquim Brandão. 2017. Deriving sonority from the structure, not the other way round: A Strict CV approach to consonant clusters. *The Linguistic Review* 34(4): 589–614.

Cavirani, Eduardo. To appear. Silent lateral actors.

Chabot, Alex. 2019. What's wrong with being a rhotic? *Glossa: A journal of general linguistics* 4(1): article 38.

SCHEER 491

Chabot, Alex. 2021. Possible and impossible languages: Naturalness, third factors, and substance-free phonology in the light of crazy rules. Doctoral dissertation, Université Côte d'Azur.

Chen, Matthew. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22(3): 129–159.

Chierchia, Gennaro. 1986. Length, syllabification and the phonological cycle in Italian. *Journal of Italian Linguistics* 8: 5–34.

Cho, Taehong. 2011. Laboratory phonology. In *The continuum companion to phonology*, ed. Bert Botma, Nancy Kula and Kuniya Nasukawa, 343–368. New York: Continuum.

Cho, Taehong, and Peter Ladefoged. 1999. Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics* 27: 207–229.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step. Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. Roger Martin, David Michaels and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1): 1–22.

Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English*. Cambridge, MA: MIT Press.

Clayton, Ian. 2010. *On the natural history of preaspirated stops*. Doctoral dissertation, University of North Carolina at Chapell Hill.

Clements, George. 1985. The geometry of phonological features. *Phonology Yearbook* 2: 225–252.

Clements, George. 1990. The role of the sonority cycle in core syllabification. In *Papers in laboratory phonology I*, ed. John Kingston and Mary Beckmann, 283–333. Cambridge: Cambridge University Press.

Clements, George. 2009. Does sonority have a phonetic basis? In *Contemporary views on architecture and representations in phonological theory*, ed. Eric Raimy and Charles Cairns, 165–175. Cambridge, MA: MIT Press.

Clements, George, and Elizabeth Hume. 1995. The internal organization of speech sounds. In *The handbook of phonological theory*, ed. John Goldsmith, 245–306. Oxford: Blackwell.

Clements, George, and Samuel Keyser. 1983. *CV phonology. A generative theory of the syllable*. Cambridge, MA: MIT Press.

Cohn, Abigail. 1998. The phonetics–phonology interface revisited: Where's phonetics? *Texas Linguistic Forum* 41: 25–40.

Collischonn, Gisela, and Cristine Costa. 2003. Resyllabification of laterals in Brazilian Portuguese. *Journal of Portuguese Linguistics* 2(2): 31–54.

Coltheart, Max. 1999. Modularity and cognition. *Trends in Cognitive Sciences* 3(3): 115–120.

Cyran, Eugeniusz, and Morgan Nilsson. 1998. The Slavic [w > v] shift: A case for phonological strength. In *Structure and interpretation. Studies in phonology*, ed. Eugeniusz Cyran, 89–100. Lublin: Pase.

D'Alessandro, Roberta, and Tobias Scheer. 2015. Modular PIC. *Linguistic Inquiry* 46(4): 593–624.

Dresher, Elan. 2014. The arch not the stones: Universal feature theory without universal features. *Nordlyd* 41: 165–181.

Dresher, Elan. 2018. Contrastive Hierarchy Theory and the nature of features. *Proceedings of the 35th West Coast Conference on Formal Linguistics* 35: 18–29.

Dressler, Wolfgang. 1981. External evidence for an abstract analysis of the German velar nasal. In *Phonology in the 1980s*, ed. Didier Goyvaerts, 445–467. Ghent: Story-Scientia.

Faust, Noam, Adèle Jatteau, and Tobias Scheer. 2018. Two phonologies. Paper presented at the 26th Manchester Phonology Meeting, Manchester, 24–26 May.

https://doi.org/10.1017/cnj.2022.22 Published online by Cambridge University Press

Fletcher, Harvey. 1972. *Speech and hearing in communication*. 2nd edition. Huntington, N.Y.: Krieger. [1929]

Fodor, Jerry. 1983. *The modularity of the mind. An essay on faculty psychology*. Cambridge, MA: MIT-Bradford.

Gerrans, Philip. 2002. Modularity reconsidered. *Language and Communication* 22(3): 259–268.

Giles, S., and K. Moll. 1975. Cinefuorographic study of selected allophones of English /l/. *Phonetica* 31: 206–227.

Goddard, Ives, Charles F. Hockett, and Karl V. Teeter. 1972. Some Errata in Bloomfield's Menomini. *International Journal of American Linguistics* 38: 1–5.

Golston, Chris. 2018. Phi-features in animal cognition. *Biolinguistics* 12: 55–98.

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

Gordon, Matthew, Edita Ghushchyan, Bradley McDonnell, Daisy Rosenblum, and Patricia Shaw. 2012. Sonority and central vowels: A cross-linguistic phonetic study. In *The sonority controversy*, ed. Steve Parker, 219–256. Berlin: de Gruyter.

Goudbeek, Martijn, Roel Smits, Anne Culter, and Daniel Swingley. 2005. Acquiring auditory and phonetic categories. In *Handbook of categorization in cognitive science*, ed. Lefebvre Claire and Henri Cohen, 497–513. Amsterdam: Elsevier.

Gouskova, Maria, and Michael Becker. 2016. Source-oriented generalizations as grammar inference in Russian vowel deletion. *Linguistic Inquiry* 47(3): 391–425.

Guasti, Theresa, and Marina Nespor. 1999. Is syntax phonology-free? In *Phrasal phonology*, ed. René Kager and Wim Zonneveld, 73–97. Nijmegen: Nijmegen University Press.

Gussmann, Edmund. 1998. Domains, relations, and the English agma. In *Structure and interpretation. Studies in phonology*, ed. Eugeniusz Cyran, 101–126. Lublin: Folium.

Hale, Ken. 1973. Deep-surface canonical disparities in relation to analysis and change: An Australian case. In *Diachronic, areal, and typological linguistics*, ed. T.A. Seboek, 401–458. The Hague: Mouton.

Hale, Mark, Madelyn Kissock, and Charles Reiss. 2007. Microvariation, variation and the features of Universal Grammar. *Lingua* 117: 645–665.

Hale, Mark, and Charles Reiss. 2000. "Substance abuse" and "dysfunctionalism": Current trends in phonology. *Linguistic Inquiry* 31(1): 157–169.

Hale, Mark, and Charles Reiss. 2008. *The phonological enterprise*. Oxford: Oxford University Press.

Hamann, Silke. 2011. The phonetics–phonology interface. In *The continuum companion to phonology*, ed. Nancy Kula, Bert Botma and Kuniya Nasukawa, 202–224. London: Continuum.

Hamann, Silke. 2014. Phonetics–phonology mismatches. Paper presented at Old World Conference in Phonology, Leiden, 22–25 January.

Hammond, Michael. 1997. Vowel Quantity and syllabification in English. *Language* 73: 1–17.

Hankamer, Jorge, and Judith Aissen. 1974. The sonority hierarchy. In *Papers from the parasession on natural phonology*, ed. A. Bruck, R. Fox and M. La Galy, 131–145. Chicago: Chicago Linguistic Society.

Hansson, Gunnar Ó. 2001. Remains of a submerged continent: Preaspiration in the languages of Northwest Europe. In *Historical Linguistics 1999*, ed. L. Brinton, 157–73. Amsterdam: Benjamins.

Hargus, Sharon. 1993. Modeling the phonology–morphology interface. In *Studies in lexical phonology*, ed. Sharon Hargus and Ellen Kaisse, 45–74. New York: Academic Press.

Harnad, Stevan. 2003. Categorical perception. In *Encyclopedia of cognitive science*, ed. Lynn Nadel. Chichester: Wiley.

Harnad, Stevan. 2005. To cognize is to categorize: Cognition is categorization. In *Handbook of categorization in cognitive science*, ed. Lefebvre Claire and Henri Cohen, 19–43. Amsterdam: Elsevier.

Harrington, Jonathan, Felicitas Kleber, and Ulrich Reubold. 2008. Compensation for coarticulation, /u/-fronting, and sound change in standard southern British: An acoustic and perceptual study. *Journal of the Acoustical Society of America* 123: 2825–2835.

Harris, John. 1990. Segmental complexity and phonological government. *Phonology* 7(2): 255–300.

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

Hayes, Bruce. 1986. Inalterability in CV Phonology. *Language* 62: 321–351.

Hayes, Bruce. 1995. *Metrical Stress Theory. Principles and Case Studies*. Chicago: University of Chicago Press.

Hayes, Bruce. 2009. *Introductory Phonology*. Oxford: Wiley-Blackwell.

Heffner, Roe-Merrill S. 1950. *General phonetics*. Madison: The University of Wisconsin Press.

Helgason, Pétur. 2002. *Preaspiration in the Nordic languages*. Doctoral Dissertation, Stockholm University.

Henton, C. G. 1983. Changes in the vowels of received pronunciation. *Journal of Phonetics* 11(4): 353–371.

Hermans, Ben, and Marc van Oostendorp. 2005. Against the sonority scale: Evidence from Frankish tones. In *Organizing grammar. Studies in Honor of Henk van Riemsdijk*, ed. Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz and Jan Koster, 206–221. Berlin: Mouton de Gruyter.

Hooper, Joan. 1976. *An introduction to Natural Generative Phonology*. New York: Academic Press.

van der Hulst, Harry. 1994. Radical CV Phonology: The locational gesture. *UCL Working Papers in Linguistics* 6: 439–478.

van der Hulst, Harry. 1995a. Radical CV Phonology: The categorial gesture. In *Frontiers of phonology: Atoms, structures, derivations*, ed. Jacques Durand and Francis Katamba, 80–116. London and New York: Longman.

van der Hulst, Harry. 1995b. Head-dependency relations in the representation of signs. In *Sign language research 1994. Proceedings of the 4th European congress on Sign language research*, ed. H. Bos and T. Schermer, 11–38. Hamburg: Signum Press.

van der Hulst, Harry. 1999. Features, segments and syllables in Radical CV Phonology. In *Phonologica 1996*, ed. John Rennison and Klaus Kühnhammer. The Hague: Holland Academic Graphics.

Hume, Elizabeth, and David Odden. 1996. Reconsidering [consonantal]. *Phonology* 13(3): 345–376.

Hyman, Larry. 2001. The limits of phonetic determinism in phonology: *NC revisited. In *The role of speech perception in phonology*, ed. Elizabeth Hume and Keith Johnson, 141–185. New York: Academic Press.

Hyman, Larry. 2006. Word-prosodic typology. *Phonology* 23(2): 225–257.

Hyman, Larry. 2011. Tone: Is it different? In *The handbook of phonological theory*, 2nd ed., ed. John Goldsmith, Jason Riggle and Alan C. L. Yu, 197–239. London: Blackwell.

Hyman, Larry M., and Russell G. Schuh. 1974. Universals of tone rules: Evidence from West Africa. *Linguistic Inquiry* 5(1): 81–115.

Inkelas, Sharon. 1995. The consequences of optimization for underspecification. *Proceedings of the North East Linguistic Society* (NELS) 25: 287–302. [ROA #40].

Inkelas, Sharon, and Draga Zec. 1990. Prosodically constrained syntax. In *The phonology–syntax connection*, ed. Sharon Inkelas and Draga Zec, 365–378. Chicago: Chicago University Press.

Iosad, Pavel. 2012. *Representation and variation in substance-free phonology. A case study in Celtic*. Doctoral dissertation, University of Tromsø.

Iosad, Pavel. 2017. *A substance-free framework for phonology. An analysis of the Breton dialect of Bothoa*. Edinburgh: Edinburgh University Press.

Iverson, Gregory K., and Joseph C. Salmons. 2011. Final devoicing and final laryngeal neutralization. In *The Blackwell companion to phonology*, ed. Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice, 1622–1643. New York: Wiley-Blackwell.

Jackendoff, Ray. 2002. *Foundations of language. brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

Kahn, Daniel. 1980. *Syllable-based generalizations in English phonology*. New York: Garland Press.

Kay, Paul, Brent Berlin, Luisa Maffi, William R. Merrifield, and Richard Cook. 2009. *The world color survey*. Stanford, Cal.: CSLI.

Kay, Paul, and Chad K. McDaniel. 1978. The linguistic significance of the meanings of basic color terms. *Language* 54: 610–646.

Kaye, Jonathan. 1990. Government in phonology: The case of Moroccan Arabic. *The Linguistic Review* 6: 131–159.

Kaye, Jonathan, and Jean Lowenstamm. 1984. De la syllabicité. In *Forme sonore du langage*, ed. François Dell, Daniel Hirst et Jean-Roger Vergnaud, 123–159. Paris: Hermann.

Kaye, Jonathan, Jean Lowenstamm, and Jean-Roger Vergnaud. 1990. Constituent structure and government in phonology. *Phonology Yearbook* 7: 193–231.

Keating, Patricia. 1983. Comments on the jaw and syllable structure. *Journal of Phonetics* 11: 401–406.

Keating, Patricia. 1985. Universal phonetics and the organization of grammars. In *Phonetic linguistics essays in Honour of Peter Ladefoged*, ed. Victoria Fromkin, 115–132. Orlando: Academic Press.

Kilbourn-Ceron, Oriana, and Morgan Sonderegger. 2018. Boundary phenomena and variability in Japanese high vowel devoicing. *Natural Language and Linguistic Theory* 36(1): 175–217.

Kingston, John. 2019. The interface between phonetics and phonology. In *The Routledge handbook of phonetics*, ed. William F. Katz and Peter F. Assmann, 359–400. Abingdon: Routledge.

Kiparsky, Paul. 1968–1973. How abstract is phonology? Manuscript circulated since 1968 and published in*: Three dimensions of linguistic theory*, ed. Osamu Fujimura, 5–56. Tokyo: TEC.

Klatt, Dennis H. 1973. Interaction between two factors that influence vowel duration. *The Journal of the Acoustical Society of America* 54: 1102–1104.

de Lacy, Paul. 2002. *The formal expression of markedness*. Doctoral dissertation, University of Massachusetts.

Ladefoged, Peter, and Sandra Ferrari-Disner. 2012. *Vowels and Consonants*. 3rd ed. Oxford: Wiley-Blackwell.

Lee-Kim, S.-I., Davidson, L., and Hwang, S. 2013. Morphological effects on the articulation of English intervocalic /l/. *Laboratory Phonology* 4(2): 475–511.

Lehiste, Ilse. 1960. *An acoustic–phonetic study of internal open juncture*. Basel: Karger (supplement to Phonetica 5).

Lombardi, Linda. 2001. Why Place and Voice are different. In *Segmental phonology in Optimality Theory: Constraints and representations*, ed. L. Lombardi, 13–45. Cambridge: Cambridge University Press.

Lowenstamm, Jean. 1991. Vocalic length and centralization in two branches of Semitic (Ethiopic and Arabic). In *Semitic studies in Honor of Wolf Leslau on the occasion of his 85th birthday*, ed. Alan S. Kaye, 949–965. Wiesbaden: Harrassowitz.

Lowenstamm, Jean. 1999. The beginning of the word. In *Phonologica 1996*, ed. John Rennison and Klaus Kühnhammer, 153–166. The Hague: Holland Academic Graphics.

Lowenstamm, Jean. 2011. The Phonological pattern of phi-features in the perfective paradigm of Moroccan Arabic. *Brill's Annual of Afroasiatic Languages and Linguistics* 3: 140–201.

Mackenzie, Sara, Erin Olson, Meghan Clayards, and Michael Wagner. 2018. North American /l/ both darkens and lightens depending on morphological constituency and segmental context. *Laboratory Phonology* 9: article 13.

Marr, David. 1982. *Vision*. San Francisco: Freeman.

McCarthy, John. 1998. Morpheme structure constraints and paradigm occultation. In *Chicago Linguistics Society 32. Part 2: The Panels*, ed. M. Catherine Gruber, Derrick Higgins, Kenneth Olson and Tamra Wysocki, 123–150. Chicago: Chicago Linguistic Society.

McCarthy, John. 2003. Sympathy, cumulativity, and the Duke-of-York Gambit. In *The syllable in Optimality Theory*, ed. Caroline Féry and Ruben van de Vijver, 23–76. Cambridge: Cambridge University Press.

McGurk, Harry, and John MacDonald. 1976. Hearing lips and seeing voices. *Nature* 264: 746–748.

Mielke, Jeff. 2008. *The emergence of distinctive features*. Oxford: Oxford University Press.

Moravcsik, Edith. 2000. Infixation. In *morphology. An international handbook on inflection and word-formation, Vol.1*, ed. Geert Booij, 545–552. Berlin: de Gruyter.

Moreton, Elliott, Gary Feng, and Jennifer Smith. 2005. Syllabification, sonority, and perception: New evidence from a language game. *Proceedings from the Annual Meeting of the Chicago Linguistic Society* 41: 341–355.

Nasukawa, Kuniya. 2005. *A unified approach to nasality and voicing*. Berlin: Mouton deGruyter.

Nasukawa, Kuniya. 2015. Recursion in the lexical structure of morphemes. In *Representing structure in phonology and syntax*, ed. Marc van Oostendorp and Henk van Riemsdijk, 211–238. Berlin: de Gruyter.

Nespor, Marina, and Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris.

Odden, David. 2006. Phonology ex nihilo, aka Radical Substance-Free Phonology and why I might recant. Paper presented at Phonological seminar, Tromsø, 6 December.

Odden, David. 2022. Radical substance-free phonology and feature learning. *Canadian Journal of Linguistics* 67(4), 500–551.

Ohala, John. 1974. Phonetic explanation in phonology. In *Papers from the parasession on Natural Phonology*, ed. Anthony Bruck, Robert Fox and William La Galy, 251–274. Chicago: Chicago Linguistic Society.

Ohala, John. 1990. There is no interface between phonology and phonetics: A personal view. *Journal of Phonetics* 18(2): 153–171.

Ohala, John. 1992. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In *Papers from the Parasession on the Syllable, Chicago Linguistic Society*, 319–338. Chicago: Chicago Linguistic Society.

Ohala, John, and Haruko Kawasaki. 1984. Prosodic phonology and phonetics. *Phonology Yearbook* 1: 113–127.

Pagliano, Claudine. 2003. *L'épenthèse consonantique en français. Ce que la syntaxe, la sémantique et la morphologie peuvent faire à la phonologie*. Thèse de doctorat, Université de Nice.

Parker, Steve. 2001. Non-optimal onsets in Chamicuro: An inventory maximised in coda position. *Phonology* 18: 361–386.

Parker, Steve. 2008. Sound level protrusions as physical correlates of sonority. *Journal of Phonetics* 36: 55–90.

Parker, Steve. 2017. Sounding out sonority. *Language and Linguistics Compass* 11(19): e12248.

Paster, Mary. 2006. *Phonological conditions on affixation*. Doctoral dissertation, University of California at Berkeley.

Perlmutter, David. 1992. Sonority and syllable structure in American Sign Language. *Linguistic Inquiry* 23(3): 407–442.

Pierrehumbert, Janet, and Mary Beckman. 1988. *Japanese tone Structure*. Cambridge: MIT Press.

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

Pöchtrager, Markus A., and Jonathan Kaye. 2013. GP2.0. *SOAS Working Papers in Linguistics and Phonetics* 16: 51–64.

Price, C. J. 1980. Sonority and syllabicity: Acoustic correlates of perception. *Phonetica* 37: 327–343.

Prince, Alan, and Paul Smolensky. 2004. *Optimality Theory. Constraint Interaction in Generative Grammar*. Oxford: Blackwell. [1993]

Rasin, Ezer. 2018. Modular interactions in Phonology. Doctoral dissertation, Massachusetts Institute of Technology.

Rasin, Ezer, and Roni Katzir. 2015. A learnability argument for constraints on underlying representations. In *Proceedings of the 45th Annual Meeting of the NorthEast Linguistic Society (NELS), Vol.2*, ed. Thuy Bui and Deniz Özyildiz, 267–288. Cambridge, MA: GLSA.

Reiss, Charles. 2003. Towards a theory of fundamental phonological relations. In *Asymmetry in phonology, Vol.2*, ed. Anna Maria Di Sciullo, 215–238. Amsterdam: Benjamins.

Reiss, Charles, and Veno Volenec, Conquer primal fear: Phonological features are innate and Substance free. In *Canadian Journal of Linguistics* 67(4), 581–610.

Rice, Keren. 1992. On deriving sonority: A structural account of sonority relationships. *Phonology* 9: 61–99.

Rizzolo, Olivier. 2002. Du leurre phonétique des voyelles moyennes en français et du divorce entre Licenciement et Licenciement pour gouverner. Thèse de doctorat, Université de Nice.

Rose, Sharon, and Peter Jenks. 2011. High tone in Moro: ffects of prosodic categories and morphological domains. *Natural Language and Linguistic Theory* 29(1): 211–250.

Sagiv, Noam, and Jamie Ward. 2006. Crossmodal interactions: Lessons from synesthesia. *Progress in Brain Research* 155: 259–271.

Samuels, Bridget. 2009. The structure of phonological theory. Doctoral dissertation, Harvard University.

Samuels, Bridget. 2012. The emergence of phonological forms. In *Towards a biolinguistic understanding of grammar. Essays on interfaces*, ed. Anna Maria Di Sciullo, 193–213. Amsterdam: Benjamins.

Samuels, Bridget, Samuel Andersson, Ollie Sayeed, and Bert Vaus. Getting ready for Primetime: Paths to acquiring substance-free phonology. In *Canadian Journal of Linguistics* 67(4):552–580.

Sandler, Wendy. 1993. A sonority cycle in American Sign Language. *Phonology* 10: 243–279.

Sandler, Wendy, and Diane Lillo-Martin. 2006. *Sign language and linguistic universals*. Cambridge: Cambridge University Press.

Saussure, Ferdinand de. 1972. *Cours de linguistique générale*. Paris: Payot. [1916]

Scheer, Tobias. 2004. *A lateral theory of phonology. Vol.1: What is CVCV, and why should it be?* Berlin: Mouton de Gruyter.

Scheer, Tobias. 2011. *A guide to morphosyntax–phonology interface theories. How extra-phonological information is treated in phonology since Trubetzkoy's Grenzsignale.* Berlin: Mouton de Gruyter.

Scheer, Tobias. 2012a. *Direct interface and one-channel translation. A non-diacritic theory of the morphosyntax–phonology interface. Vol.2 of A Lateral Theory of phonology.* Berlin: de Gruyter.

Scheer, Tobias. 2012b. Melody-free syntax and two phonologies. Paper presented at the annual conference of the Réseau Français de Phonologie (RFP), Paris, 25–27 June.

Scheer, Tobias. 2014a. The initial CV: Herald of a non-diacritic interface theory. In *The Form of structure, the structure of form. Essays in Honor of Jean Lowenstamm*, ed. Sabrina Bendjaballah, Noam Faust, Mohamed Lahrouchi and Nicola Lampitelli, 315–330. Amsterdam: Benjamins.

Scheer, Tobias. 2014b. Spell-out, post-phonological. In *Crossing phonetics–phonology lines*, ed. Eugeniusz Cyran and Jolanta Szpyra-Kozlowska, 255–275. Newcastle upon Tyne: Cambridge Scholars.

Scheer, Tobias. 2015. How diachronic is synchronic grammar? Crazy rules, regularity and naturalness. In *The Oxford handbook of historical phonology*, ed. Patrick Honeybone and Joseph C. Salmons, 313–336. Oxford: Oxford University Press.

Scheer, Tobias. 2016. Melody-free syntax and phonologically conditioned allomorphy. *Morphology* 26(4): 341–378.

Scheer, Tobias. 2017. Voice-induced vowel lengthening. *Papers in historical phonology* 2: 116–151. <http://journals.ed.ac.uk/pihph/issue/view/150>

Scheer, Tobias. 2019a. Phonetic arbitrariness: A cartography. *Phonological Studies* 22: 105–118.

Scheer, Tobias. 2019b. Sonority is different. *Studies in Polish Linguistics* 14 (special volume 1): 127–151.

Scheer, Tobias. 2020. Final devoicing is not phonological. Paper presented at the Old World Conference in Phonology (OCP 17), Warsaw, 5–7 February.

Scheer, Tobias, and Nancy C. Kula. 2018. Government Phonology: Element theory, conceptual issues and introduction. In *The Routledge Handbook of Phonological Theory*, ed. S.J. Hannahs and Anna Bosch, 226–261. Oxford: Routledge.

Schwartz, Geoff. 2013. A representational parameter for onsetless syllables. *Journal of Linguistics* 49(3): 613–646.

Schwartz, Geoff. 2017. Formalizing modulation and the emergence of phonological heads. *Glossa: A journal of general linguistics* 2(1), article 81.

Segal, Gabriel. 1996. The modularity of theory of mind. In *Theories of theories of mind*, ed. P. Carruthers and P. Smith, 141–157. Cambridge: Cambridge University Press.

Ségéral, Philippe. 1996. L'apophonie en ge'ez. In *Studies in Afroasiatic Grammar*, ed. Jacqueline Lecarme, Jean Lowenstamm and Ur Shlonsky, 360–391. The Hague: Holland Academic Graphics.

Ségéral, Philippe, and Tobias Scheer. 2001a. La Coda-Miroir. *Bulletin de la Société de Linguistique de Paris* 96: 107–152.

Ségéral, Philippe, and Tobias Scheer. 2001b. Abstractness in phonology: The case of virtual geminates. In *Constraints and preferences*, ed. Katarzyna Dziubalska-Kołaczyk, 311–337. Berlin: Mouton de Gruyter.

Ségéral, Philippe, and Tobias Scheer. 2008a. Positional factors in lenition and fortition. In *Lenition and fortition*, ed. Joaquim Brandão de Carvalho, Tobias Scheer and Philippe Ségéral, 131–172. Berlin: Mouton de Gruyter.

Ségéral, Philippe, and Tobias Scheer. 2008b. The Coda Mirror, stress and positional parameters. In *Lenition and fortition*, ed. Joaquim Brandão de Carvalho, Tobias Scheer and Philippe Ségéral, 483–518. Berlin: Mouton de Gruyter.

Seinhorst, Klaas, Paul Boersma, and Silke Hamann. 2019. Iterated distributional and lexicon-driven learning in asymmetric neural network explains the emergence of features and dispersion. In *Proceedings of the 19th International Congress of Phonetic Sciences*, ed. S. Calhoun, P. Escudero, M. abain and P. Warren, 1134–1138. Canberra: Australasian Speech Science and Technology Association Inc.

Selkirk, Elisabeth. 1981. On prosodic structure and its relation to syntactic structure. In *Nordic prosody II*, ed. Thorstein Fretheim, 111–140. Trondheim: TAPIR. [1978]

Shannon, Claude E. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27: 379–423.

Smith, Jennifer L., and Elliott Moreton. 2012. Sonority variation in stochastic Optimality Theory: Implications for markedness hierarchies. In *The Sonority Controversy*, ed. Steve Parker, 167–194. Berlin: de Gruyter.

Solé, Maria-Josep, Larry Hyman, and Kemmonye C. Monaka. 2010. More on post-nasal devoicing: The case of Shekgalagari. *Journal of Phonetics* 38(4): 604–615.

Solé, Maria-Josep, Ronald Sprouse, and John Ohala. 2008. Voicing control and nasalization. *Laboratory Phonology* 11: 127–128.

Steriade, Donca. 1982. Greek prosodies and the nature of syllabification. Doctoral dissertation, Massachusetts Institute of Technology.

Steriade, Donca. 1994. *Positional neutralization and the expression of contrast*. Ms, University of California, Los Angeles.

Steriade, Donca. 1999. Alternatives to syllable-based accounts of consonantal phonotactics. In *Proceedings of LP'98: Item Order in Language and Speech, Vol. 1*, ed. Osamu Fujimura, Brian D. Joseph and Bohumil Palek, 205–245. Prague: Karolinum Press.

Stevens, Kent A. 2012. The vision of David Marr. *Perception* 41: 1061–1072.

Stilp, Christian E., and Keith R. Kluender. 2010. Cochlea-scaled entropy, not consonants, vowels, or time, best predicts speech intelligibility. *PNAS* 107: 12387–12392.

Stokoe, William C. 1978. *Sign language structure*. Silver Spring, MD: Linstok Press. [1960]

Strycharczuk, Patrycja, and James M. Scobbie. 2016. Gradual or abrupt? The phonetic path to morphologisation. *Journal of Phonetics* 59: 76–91.

Szendrői, Kriszta. 2003. A stress-based approach to the syntax of Hungarian focus. *The Linguistic Review* 20(1): 37–78.

Szendrői, Kriszta. 2004. A stress-based approach to climbing. In *Verb clusters. A study of Hungarian, German and Dutch*, ed. Katalin É. Kiss and Henk van Riemsdijk, 205–233. Amsterdam: Benjamins.

Szigetvári, Péter. 2008a. What and where? In *Lenition and fortition*, ed. Joaquim Brandão de Carvalho, Tobias Scheer and Philippe Ségéral, 93–129. Berlin: Mouton de Gruyter.

Szigetvári, Péter. 2008b. Two directions for Lenition. In *Lenition and fortition*, ed. Joaquim Brandão de Carvalho, Tobias Scheer and Philippe Ségéral, 561–592. Berlin: Mouton de Gruyter.

Szigetvári, Péter, and Tobias Scheer. 2005. Unified representations for the syllable and stress. *Phonology* 22(1): 37–75.

Tamminga, Meredith. 2018. Modulation of the following segment effect on English coronal stop deletion by syntactic boundaries. *Glossa: A journal of general linguistics* 3(1): 86.

Tanner, James, Morgan Sonderegger, and Michael Wagner. 2017. Production planning and coronal stop deletion in spontaneous speech. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 8: 15.

Turton, Danielle. 2017. Categorical or gradient? An ultrasound investigation of /l/-darkening and vocalization in varieties of English. *Laboratory Phonology* 8: article 13.

Vaux, Bert. 2003. Syllabification in Armenian, Universal Grammar, and the lexicon. *Linguistic Inquiry* 34(1): 91–125.

Vaux, Bert. 2005. Formal and empirical arguments for Morpheme Structure Constraints. Paper presented at Linguistic Society of America, Oakland, January 5th.

Vaux, Bert, and Bridget Samuels. 2018. Abstract underlying representations in prosodic structure. In *Shaping phonology*, ed. Diane Brentari and Jackson Lee, 146–181. Chicago: University of Chicago Press.

Vennemann, Theo. 1972. Sound change and markedness theory: On the history of the German consonant system. In *Linguistic change and generative theory. Essays from the UCLA Conference on historical linguistics in the perspective of transformational theory (1969)*, ed. R.P. Stockwell and R.K.S. Macaulay, 230–274. Bloomington: Indiana Univ. Press.

Volenec, Veno, and Charles Reiss. 2018. Cognitive phonetics: The transduction of distinctive features at the phonology–phonetics interface. *Biolinguistics* 11: 251–294.

Wagner, Michael. 2012. Locality in phonology and production planning. *McGill Working Papers in Linguistics* 22: 1–18.

Wilson, Stephen. 1986. Metrical structure in Wakashan phonology. In *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*, ed. Vassiliki Nikiforidou, Mary Van Clay, Mary Niepokuj and Deborah Feder, 283–291. Berkeley: Berkeley Linguistics Society.

Wright, Richard. 2004. A review of perceptual cues and cue robustness. In *Phonetically Based Phonology*, ed. Bruce Hayes, Donca Steriade and Robert Kirchner, 34–57. Cambridge: Cambridge University Press.

Yip, Moira. 1996. Lexicon optimization in languages without alternations. In *Current trends in phonology. Models and methods, Vol.2*, ed. Jacques Durand and Bernard Laks, 759–790. Salford, Manchester: ESRI.

Yu, Alan C. L. 2007. *A natural history of infixation*. Oxford: Oxford University Press.

Zec, Draga. 1995. Sonority constraints on syllable structure. *Phonology* 12: 85–129.