An x-bar theory of Government Phonology*

John R. Rennison and Friedrich Neubarth

Introduction

This article originated as an outline of a few revisions to standard Government Phonology, but it has now grown to define a quite comprehensive theory and model of phonology, based on previous work in Government Phonology and on insights gained through failures of the standard theory to naturally explain some of the data.

When the model of Government Phonology was set out during the 1980's, the main focus was on how to represent vowels and the structural relationships between them (proper government) and between and within other constituents like "onsets" and "codas". Certain phenomena were acknowledged as rather mysterious (e.g., the "magic licensing" of Kaye [1991/1992] 1996a), while others were rarely recognised as a problem (e.g., domain-final codas with more than two consonants). The representation of consonants always lagged behind vowels, even more so the representation of inter-consonantal relations.[1] Indirectly related to this, in the past few years a vivid discussion has arisen about higher structuring, i.e., syllable structure, which one recent school of thought considers to be quite primitive (onset-nucleus/consonant-vowel pairs — cf. Lowenstamm 1996, Szigetvári 1999), rather than mirroring the traditional view of syllabification, with its structure of onsets, nuclei, and codas (cf. Kiparsky 1981, Kaye, Lowenstamm and Vergnaud 1989, and for critical discussion see Ploch, this volume).

What we want to do in this paper is: agreeing with the view that syllable structure consists only of CV pairs, to establish this minimal structure as the basic *constituent* of phonological theory. In particular, all licensing relations are defined solely on this level. For consonants, we will present a radically minimal system of elements, which overcomes some redundancies and stipulative constraints of previous accounts and comprises a fundamentally more restrictive inventory.

Finally, we will reconsider the mechanisms of licensing (uninterpreted vowels rendering consonant clusters), arguing that one class of consonant clusters can easily be accounted for if proper government (henceforth V-government) is complemented with a second notion 'C-government' (which in part comprises interonset licensing and coda licensing, but differs in core technical details). The other class of consonant clusters, which coincides largely with the traditional branching onsets of standard Government Phonology, are analysed as complex single consonants.

Before outlining our model, we will address a question which has often been posed in recent years: how phonetic is phonology? Here, on a par with most exponents of Government Phonology, we take a very strict point of view. The framework we present here is a formal theory designed to account for the patterning of phonological data, and it is neither motivated by, nor always directly relatable to, phonetics. This is a conceptual rather than an empirical matter. Phonetic output has no resemblance to or impact on phonology. We do not want to preclude that our view on how phonological structure is represented might have interesting consequences for a different conception of phonetics as a direct interpretation (articulatory or acoustic) of these structures. But we definitely want to exclude any structural relation between phonetics and phonology, or any structural interface whatsoever.

Our present venture, perhaps better regarded as a research program, reflects what we believe to be a challenging recapitulation of the hard-gained insights from previous work by many people, combined with our own ideas about a unified model of phonology. As Jonathan Kaye instructed us, it is more fruitful to have a very restrictive theory which gets the basics right, and to stand with our backs to the wall and await objections.

1. Structure

1.1. Sylls

We assume the standard level of skeletal points (×) as the melodic side of the interface between (hierarchical) phonological structure and melody.

Throughout the whole of (linguistic, and therefore *a fortiori*) phonological structure, heads project, non-heads do not (Jackendoff 1977). We claim that phonological structure is universally right headed at the base level (although this does not hold for higher levels of metrical structure).[2] In terms of the skeleton, nuclei project, onsets do not. Thus, phonological structure is minimal: only right-headed binary branching constituents are permitted. Each skeletal point (×) is either the head or non-head of such a constituent, which we will call a *syll* (noted here with the symbol $\overline{\times}$).[3] A syll thus consists of a pair of skeletal points, the second of which is the head. The non-head position is what is traditionally labelled a consonant (or onset), the head position a vowel (or nucleus).[4] Thus, the labels C (or O) and V (or N) as well as the structure CV (ON), as the core of phonological structure, are derived from the familiar structural principles of binary branching, headedness and projection, as can be seen by comparing (1a) with its equivalent in "CVCV"-Government Phonology, (1b).

(1) a. A syll constituent

$$\begin{array}{ccc} & \overline{\times} & \\ \diagup & & \diagdown \\ \times & & \underline{\times} \end{array}$$

b. A CV (ON) pair

$$\begin{array}{cc} C & V \\ | & | \\ \times & \times \end{array}$$

(The arrow indicates inter-constituent government.)

In the present theory there is no onset constituent on any level higher than the skeleton. This means that onsets never "see each other" directly – i.e., interonset relations of every kind are mediated by the syll constituent in which the onsets are located, and therefore often subject to (even multiple) parametric constraints.

The $\overline{\times}$ node of the syll constituent is the structural side of the interface between structure and melody, and the two × nodes are the melodic side. The role of the skeleton is fundamentally diminished in this account, since the skeleton is entirely structurally determined. Elements (representing melody) can be attached to either of the two skeletal positions of a syll, but are never directly attached to the $\overline{\times}$ node. Interconstituent relations and higher-level (i.e., prosodic) phonological relations, on the other hand, can only relate to the $\overline{\times}$ node of a syll, and never directly to the skeleton. This has several benefits:

- Onsets can never relate directly to prosodic structure. So, for example, no language can require that a particular nucleus must have a filled onset, an empty onset, or an onset linked to a specific melody element, except via constraints on syll constituents.
- Nuclei, on the other hand, being heads, *can* be required by higher-level mechanisms to be filled, empty, or to have a particular melodic content.
- There is no longer any direct relationship between a nucleus and its onset of the kind expressed by interconstituent government.
- The syll, which is the minimal phonetically pronounceable phonological expression and (we claim) minimal cognitive unit is now the central structural unit of phonology.[5]

1.2. Phonological domains

At the lowest level of structure, a phonological domain consists of a sequence of one or more sylls. It is typically, but not necessarily, a phonological word. Depending on parameter settings, a phonological domain on the lowest level may or may not contain certain morphological boundaries. All licensing relations are bounded by a phonological domain, and at the edges of a domain there may be special licensing mechanisms (e.g., final empty nucleus licensing). In principle, we agree with Kaye [1993](1995) as far as the theoretical status of domains is concerned; however, as will come clear in the following, our definitions of phonological domains deviate from his and can be seen as more restrictive.

At higher structural levels (metrical, prosodic), a domain consists of a sequence of one or more lower-level domains. It is more convenient to graphically represent higher-level domains such as feet or stress-groups as a tree structure. Unless otherwise stated, in this article the term "domain" refers to a domain of the lowest level.

Licensing conditions are checked cyclically within phonological domains. But also prosodic properties like stress assignment are computed relative to phonological domains — which means that every phonological domain must contain at least one realised nucleus. A phonological domain may itself contain phonological subdomains.[6]

2. Melody

2.1. Primitives

The set of melodic primitives must be finite, universal and minimal. In this version of Government Phonology it consists of the elements *I*, *U*, *R*, *H*, *L*, *F* (where *F* is functional element, previously represented as '_' or 'e'). The interpretation of elements depends in part on whether they are attached to a consonant or vowel position. This appears to be a quite straightforward move; however, it is very important to note that the labels "C" and "V" do not have any (melodic) content on their own. They are just structural properties derived by the principles of X-bar structure. Notice further that this is the only way to implement a difference between vowels and consonants without either reserving different elements or (maybe covertly) giving "C" and "V" element status. Their most common phonetic correlates are given in the table in (2), with the usual caveat that elements are hard-wired cognitive entities, and therefore neither articulatory nor acoustic in nature. Positions in (2) with a star (*) are impossible. ("Operator" is synonymous with "non-head". "ME" stands for "melodic expression".)

(2)

Element	ME	As head of ME	As operator in ME
F	C	*stop*	*fricative/spirant*
	V	*"A" (non-high)*	*ATR*
I	C	*i-glide*	*palatal*
	V	*"I" (front)*	*front*
U	C	*u-glide*	*labial, "dark"*
	V	*"U" (rounded)*	*rounded*
R	C	*liquid*	*coronal*
	V	** (see §2.5)*	** (see §2.5)*
H	C	*fricative*	*aspiration*
	V	** (see §2.5)*	*high tone*
L	C	*nasal*	*voiced*
	V	** (see §2.5)*	*nasal/low tone*

The main difference vis-à-vis previous proposals about the set of elements is the functional element *F*, which has a different interpretation in each of the given contexts. It replaces the former elements *A*, *ʔ*, *ɨ* (ATR) and *h* (cf. Kaye, Lowenstamm, and Vergnaud 1985; Harris 1990, 1994). However, this small move removes two redundancies: the *A* and the *ɨ*-element generally occurred only in vowels (although there are different conceptions about the *A*-element, e.g., the view that the *A*- and *R*-elements can be fused into one new *A*-element – cf. Kaye 1991; Ploch 1993; Williams 1998). The stop and the noise element are exclusively consonantal.[7] These restrictions are now combined in the definition of a single element.

The retention of *R* as an element and the interpretation of *L*-head in consonants and *L*-operator in vowels as 'nasality' are an integral part of the new setup of elements. Notice that an *F*-head in consonants is interpreted as "stop" (mainly correlated with silence) and not as "occlusive", which is a misguiding term from articulatory phonetics and quite misplaced in a phonological theory. The occlusion found in nasals therefore requires no extra specification beyond an *L*-head: nasals no longer have any equivalent of the *ʔ*-element.

2.2. *Attachment/association*

In general, any element may attach to any skeletal position (be it a head or non-head ×), though with a few restrictions outlined below. An attachment may be present in the lexicon, or it may result from licensing conditions (either universal or language-specific). We reject the Obligatory Contour Principle as a part of linguistic theory; even as a heuristic it is problematic. So there is no prohibition in principle against sequences of identical elements in consecutive positions.

Elements which are not attached to a skeletal point in the lexicon (termed *floating elements*) may have some degree of leeway as to where they attach,[8] and may also (parametrically) be realised phonetically as the second part of a contour segment[9] (cf. Rennison 1998).

Multiple lexical association of melodic expressions to skeletal positions is still an open question. If permitted, multiple associations could hardly be restricted in any principled way. We will therefore continue on the working assumption that multiple associations do not exist in the lexicon — neither from entire melodic expressions nor from individual elements to more than one skeletal position. This means that long vowels and geminates do not "share" melodic expressions or individual elements; instead, either the melodies in question are present twice, or one of the positions is lexically empty and is phonetically identified by mechanisms operating at the $\bar{\times}$ level, as outlined in §4 below.

2.3. *Melodic expressions*

The set of elements which are attached to a particular skeletal point are termed a *melodic expression*.[10] In general, only a single occurrence of a non-functional element is possible within a melodic expression. Given the standard interpretation of phonological elements as unique acoustic signatures, the double or multiple occurrence of an element (with the exception of the functional element, *F*) would not change the phonetic interpretation of the expression in which it occurs. The functional element, *F*, on the other hand, can occur both as the head and as an operator of the same melodic expression. We know of no evidence which would point to multiple occurrences of the other elements (apart from the lazy elements discussed in §2.7 below), but there is ample evidence from neutral vowels in vowel harmony processes that a putative "additional" *I*- or *U*-element has no phonetic effect if that element is already present in the vowel.

Graphically, melodic expressions are represented as follows: The head element(s) come first (and are redundantly underlined). Elements following the comma are operators. The linear order of operators is in principle irrelevant. When a melodic expression has more than one element in its head, the operators that modify only the first head element have a '–' suffix, and those which do not modify the first head element ("lazy operators") a '+' suffix (see §2.7). Where it is optically convenient, the entire melodic expression is surrounded by parentheses.

2.4. Headedness

Every melodic expression that is not empty must have at least one melodic head. This runs counter to approaches that employ headedness as a parametric option to represent phonological distinctions, e.g., Cobb (1993, 1995, 1997); Charette (1994); Kaye (1994).

As can be seen from the table in (2) above, the phonetic realisation of a particular element as the head or an operator may differ. The functional element is the only element which is allowed to occur *both* as a head and an operator in the same melodic expression. This is due to its special status: the functional element has no unique acoustic signature. Instead, being functional, *F* maximises or distracts from the most discernible acoustic pattern of the respective syll position to which it is attached. In the case of consonants, maximisation means silence (stops) and distraction means noise (spirants/fricatives). With vowels maximisation is the concentration of energy in the middle of the relevant frequency band (which spans from ca. 350Hz to 2500Hz), i.e., old *A*, as opposed to *I* (significantly more energy in the upper part) and *U* (more energy in the lower part). However, as an operator, *F* distracts from the typical formant configurations (tenseness/ATR). On the acoustic signatures of the vocalic elements A, I and U cf. Harris and Lindsey (2000).

2.5. Distributional asymmetries of elements

Contrary to the past tenet of Government Phonology, elements are neither equal in strength (see §2.8 below) nor balanced in distribution. Some of the most robust asymmetries have been build into the present theory by the introduction of the *F*-element and the reanalysis of nasals. We hope that eventually all asymmetries will be derivable from more general principles (possibly in conjunction with the affinities of element pairs); for the moment, we can only observe that they seem to exist. We also do not claim that the restrictions outlined here are the sole co-occurrence restrictions on elements within a single melodic expression, but they seem to be the most frequently observed ones. Future research may reveal others.

1. Vowels can only be headed by *F*, *I* or *U*:

   Whilst melodic expressions in non-head positions (consonants) can include any selection of the six elements, either as melodic head or operator, skeletal head positions (vowels) must have either *F*, *I* or *U* as their melodic head.[11] In other words, *R*, *H* and *L* are excluded from the melodic head position of vowels. This is reflected in the segment inventories of languages: the number of consonants is usually larger than the number of vowels. There exist Circassian languages with only a single vowel (Job 1981), and a very large number of languages has precisely the three-vowel set ($\underline{\mathrm{F}}$), ($\underline{\mathrm{I}}$), ($\underline{\mathrm{U}}$) (e.g., most Australian languages — see Dixon 1980). Yet there is no language without at least a handful of consonants e.g., Pirahã, with three vowels but at least seven consonants — see Everett (1986).

2. *R* can only be attached to a skeletal non-head (onset):

   The distribution of the element *R* is even more restricted: it simply cannot be lexically attached to a nucleus. Rhotacised vowels and "syllabic" coronals in our view always involve the spreading of *R* from an onset to a nuclear position.[12,]

3. *R* backgrounds *U*:

   If *R* is present in the representation of a melodic expression (as in /ʃ/, /θ/ or /l/), an *U*-element in the same melodic expression will never encode place (labial) but rather "darkness" or laterality.

4. Operator *F* requires an $\underline{F}$ melodic head in a consonant:

   The functional element, *F*, can only occur as an operator in a consonant if the same element, $\underline{F}$, is the head of the melodic expression.

2.6. Complexity restrictions

A language may have parametric restrictions on the number or combinability of elements associated with a skeletal position. Although these restrictions are ultimately idiosyncratic, the phonological system must be able to express them in a direct way. At a particular type of skeletal position (head or non-head), there can be restrictions on the total number of elements, on the relative strength of elements, and on melodic headedness relationships.

Let us first consider the most restrictive cases. In Circassian languages such as Kabardian (Job 1981 quoting Kuipers 1960), perhaps less obviously, also in Mandarin Chinese (Kaye 2000, Neubarth and Rennison 2000), a nucleus may contain either ($\underline{F}$) or nothing; in most Australian languages (e.g., Nyangumarda — see Hoard and O'Grady 1976) a nucleus may have only a single element, or nothing (resulting in a harmonised vowel — see Rennison 1987). Given the ban on

H, *L*, *R* as melodic heads in head-× positions, only the elements *F*, *I* and *U* are available, giving the vowel set /a, i, u/.

Once a language permits more than one element in a head-× position, it seems that restrictions of other kinds come into play. In view of the well-known, affinities and non-affinities (assimilations and blocking effects) of elements sharing a tier,[13] such as *I–U*, *H–L* and *A–ATR* (now *F–F*), this is hardly surprising. Subsets of the whole set of possible elements in vowels occur in many languages, the precise choice being determined by the affinity of elements (which, in turn, enhances parsability): lexical melodic expressions with two elements from the same tier are more marked. The above mentioned affinities of elements seem to correlate with their strength values in consonants (see §2.8 below). We think that this is no coincidence. Languages with a 5-vowel system /a, e, i, o, u/ allow a maximum of two elements from the three most easily available (i.e., *F*, *I*, *U*), but with the additional restriction that a two-element melody may not contain elements of equal strength (*I* and *U*).

2.7. Contour segments as lazy elements or complex heads

Lazy elements as operators, and heads containing more than one element, are responsible for contour segments, both in onsets and in nuclei. Lazy elements and second (or later) elements of a head are realised phonetically later than all the other elements in the melodic expression. Space restrictions do not allow us to go into detail here, but for a first approximation, see Rennison (1998). Many of the traditional "branching onsets" are reanalysed in the present theory as contour segments — in particular those involving obstruent + glide or obstruent + liquid sequences. Short diphthongs and affricates are also contour segments (i.e., involve lazy elements). We represent lazy operators with the '+'-suffix and complex heads simply as an intrinsically ordered set of elements. Thus for example Austrian German [b̥^{l}] as in [b̥^{l}Ä^{U}] *blau* 'blue' is notationally represented as (FR,UI+), which can be resolved into a combination of two melodic expressions (F,U / R,UI). Labialized and palatalized consonants have an additional U or

I-element in the head. Affricates either involve a lazy F-operator or a complex head with an H-element.

2.8. Some representations of segments

At this point it seems appropriate to make explicit our assumptions about the representations of some of the more common segments. We definitely do not exclude the possibility of other phonological representations of segments; on the contrary, we expect that all logically possible licit combinations of elements should occur in some language or other, although some of them will be difficult to find, since each element in a melodic expression adds to its markedness (and therefore rarity).
In these examples, we will restrict ourselves to segments of English, French and Austrian German. Let us consider consonants first.

- Stops have an *F*-element as head (with exception of the glottal stop, which we assume to be the realisation of the empty melodic expression – cf. also Ploch 1999). The distinction of place (labial, coronal, palatal, velar) is encoded by the operators *U*, *R*, *I* and "nothing", respectively. A fortis/lenis contrast can be achieved either by adding an *H*-operator (indicating aspiration) to the fortis stop or by adding an *L*-operator (indicating voicing) to the lenis, as is commonly assumed within Government Phonology.[14]
- Fricatives either involve an F-operator combined with an F-head or an H–head. Which of the two option is appropriate for a given phonological entity is not easy to determine. When fricatives follow voicing contrasts involving an H-operator (as is generally assumed for English), one is inclined to analyse those fricatives with an F–head on a par with stops.
- Nasals are generated with an L-head, liquids and glides with an R-head or an I/U-head respectively. Notice that melodic expressions containing only one element as the head and no operators ([F]=/g/, [H]=/h/, [L]=/9/, [R]=/r/ and the glides [I]=/j/, [U]=/w/)

are always the weakest of their family, so we expect them to undergo changes most easily or to be subject to special licensing restrictions. In fact this prediction seems to be borne out. Velar stops and fricatives are the most eager to disappear, palatalize or assimilate; velar nasals are prohibited from initial onsets in the majority of languages, /r/ in traditional coda position vocalizes in English and Southern German and, finally, glides sometimes lead a double life between consonant and vowel position (e.g., give rise to diphthongs, etc.).

- Vowels in this new representation do not differ too much from traditional representations of Government Phonology, except that the *A*-element is encoded as an *F*-head (and must then always be head) and that differences in height or ATR among the mid-vowels (e.g., /ɔ/ *versus* /o/) can no longer be expressed by switching or demoting the head but only by an *F*-operator expressing ATR.
- Complex melodic expressions, i.e., contour segments, involve either complex heads or lazy operators, as outlined above.

(3) Contour segments:

Contour type		Melody (relevant part only)
Complex onset	$C_{(obstr.)}{}^{r}$	(<u>FR</u>,..)
	$C_{(obstr.)}{}^{l}$	(<u>FR</u>,..U+I+)
Palatalised	C^{j}	(..<u>I</u>)
Labialised	C^{w}	(..<u>U</u>)
Affricate	$C_{(stop \rightarrow fricative)}$	(..,F+) or (<u>FH</u>,..)
Diphthong	/aI/	(<u>F</u>,I+)
	/aU/	(<u>F</u>,U+)
	/0I/	(<u>F</u>,U–I+)

2.9. The melodic strength of elements

The six phonological elements are not of equal melodic strength and type. The preliminary strength metric of Rennison (1998) is here redefined in order to achieve greater systematicity. It should be noted that the given values are derived heuristically from empirical considerations rather than grounded on theoretical principles. However, we believe that this move away from calculating complexity only by the number of elements (cf. Harris 1994) to a more complex balance is a positive and necessary one, if we are to recapture and exploit what was lost when charming theory was abandoned (cf. Kaye, Lowenstam and Vergnaud 1985).

The simple rationale behind the system is that heads decrease the strength of a consonant, whereas operators add to it. The first guideline encodes what has been intuitively captured by some versions of sonority hierarchy, the second formalizes complexity.

(4) The strength of elements (for governing relations)

Element	F	I	U	H	L	R
Value as head	0	–5	–5	–2	–4	–6
Value as operator	–2	+1	+1	+3	+3	+4

(5) Some examples:

		ME		heads	operators	total
/g/	=	($\underline{\text{F}}$)	=	0		= 0
/t/	=	($\underline{\text{F}}$,RH)	=	0	+4+3	= 7
/s/	=	($\underline{\text{H}}$,R)	=	–2	+4	= 2
/b/	=	($\underline{\text{F}}$,U)	=	0	+1	= 1
/m/	=	($\underline{\text{L}}$,U)	=	–4	+1	= –3
/l/	=	($\underline{\text{R}}$,UI)	=	–6	+1+1	= –4
/j/	=	($\underline{\text{I}}$)	=	–5		= –5
/r/	=	($\underline{\text{R}}$)	=	–6		= –6

Since a consonant is a non-head, the strength calculated for its melodic expression percolates up to its dominating $\overline{\times}$ node and there represents the consonantal strength of the syll. We will indicate this as a superscript numerical value on $\overline{\times}$. A syll with an unlicensed, hence phonetically interpreted nucleus will have a "V" subscript, so that all information relevant to government-related licensing (see the next section) is represented at the $\overline{\times}$ level.

3. Licensing

In general, melody is always licensed (and therefore phonetically fully realised), unless it can be (parametrically) suppressed by government.[15] Thus, in the normal case, the melody which is lexically associated with a skeletal position will be realised phonetically at that position. On the other hand, a position which has no lexical melody[16] will be *identified* (realised) phonetically unless it is licensed to remain uninterpreted. In other words, every skeletal position has inherent phonetic content, even if it is melodically "empty".

There are several ways of licensing phonological structures, all of which are subject to parametric variation as to whether or not they apply. One licensing mechanism is classic proper government, henceforth *V-government*. What we will also define here is a second notion of government, namely *C-government*, which replaces the interonset licensing mechanisms outlined by various people in the past years.

All government mechanisms operate exclusively from right to left and at the syll level.

(6) Government

A syll σ_i can govern its left neighbour σ_{i-1} iff at least one of the two following mechanisms is parametrically enabled and obtains:

1. the V of σ_i is unlicensed (=V-government), or;
2. the consonantal melodic strength of σ_i is greater than that of σ_{i-1} (=C-government).

(7) Licensing
The head of a syll σ is licensed to remain phonetically uninterpreted iff at least one of the three following conditions is parametrically enabled and obtains:
1. σ is governed (by C or V-government);
2. σ has a melodically empty nucleus and is at the right edge of the phonological domain;
3. σ forms a geminate structure with its right neighbour.

3.1. Licensing by government

The empty head of an unlicensed syll will always be phonetically interpreted, but licensed positions are not so straightforward. The degree of variation is wide, covering the optional non-suppression of "empty" nuclei in Koromfe, the total suppression of lexical melody in Odawa, the partial suppression of lexical melody in unstressed syllables in Russian and the non-suppression of lexical melody in German. In English, governed lexical melodies are realised with the phonetic content of an ungoverned "empty" nucleus (i.e., usually [7]), as in *photograph* [f'7Ut7gr,ä:f] – *photographer* [f7t'0gr7f7] – *photographic* [f,7Ut7gr'æfIk].

The nucleus of a lexically empty syll which is not governed must be phonetically identified. These are the classical "filled empty nuclei" which are realised with a variety of vowels in various languages: [ɨ] in Arabic, [7] in English, French and German. Such sylls are also susceptible to phonetic identification by other means (if such are parametrically available in the language), such as vowel harmony or the formation of "syllabic" consonants.

If the head of a licensed syll has no lexical melody, it is permitted to remain phonetically unidentified (i.e., silent). Language-specifically (parametrically), such positions (the classic "properly governed empty nuclei" and licensed final empty nuclei) can receive the weakest form of phonetic filling (e.g., schwa in many Gur languages such as Koromfe and Mòoré). However, in cases of parametrical phonetic expression of licensed empty nuclei, the phonetic expression of *unlicensed* nuclei must be different (i.e., melodically stronger).

3.2. Licensing by the final empty nucleus parameter

According to our definition of "final empty nuclei", given in (7b) above, there are two possibilities for an empty nucleus of a domain-final syll. If the final empty nucleus parameter is not activated, this nucleus must be phonetically interpreted (either by assimilation/harmony or with a default, usually schwa-like vowel).[17] However, in languages where the final empty nucleus parameter is operative, a domain-final nucleus is phonetically silent. In such cases, the final empty nucleus behaves as if it were governed, even though no governor is available.

To our knowledge, no previous theory has provided a principled account of the many languages which allow more sonorous word-final consonants, but ban final obstruents. We propose that this is due to the final empty nucleus parameter setting "C-governed", where the virtual governor has default strength, i.e., zero in our system.

(8) Parametrically, languages can choose between treating a domain-final melodically empty nucleus as being virtually V-governed or C-governed.

Virtual V-government is the classic Final Empty Nucleus of Kaye (1990). Virtual C-government has not previously been proposed, but is clearly needed for languages which allow only the more sonorant consonants in final positions, e.g. Koromfe (only /l/ and nasals) and Fulfulde (only /l/, /r/, nasals and glides).

Parallel to the final empty nucleus parameter, we observe that languages which allow word-initial *sC* sequences or other consonant sequences that are not analysable as complex melodic expressions must have some device which regulates the occurrence of silent nuclei at the left edge of a phonological domain. It seems as if it were the mirror image of licensing a final empty nucleus, i.e., a demand for governing capabilities. We therefore stipulate the First Syll Requirement parameter given in (9).

(9) First Syll Requirement parameter
Parametrically, languages can choose which licensing capacities must be present in the first non-empty syll of a phonological domain
a. V–government only, or
b. V or C–government

Option (9a) requires the first nucleus of a phonological domain to be unlicensed, hence phonetically realised. This permits only #CV. Option (9b) permits initial consonant clusters (i.e. $\#C_1C_2V$), but with the absolute requirement that the melodic strength of C_1 is greater than 0, thus excluding nasals glides, liquids, non-coronal fricatives, unmarked (by *H* or *L*) velar stops/spirants from the first position of such a cluster.

This is one of the very marked parameter settings responsible for initial *sC* clusters in some Indo-European languages, where the (second) C-position may or may not be a complex onset. For more detailed discussion, see §5.2 below.

3.3. Geminate licensing

That geminates form a licensing domain independent of government is not absolutely straightforward. However, we want to argue that in geminates it is the first consonant which is lexical and the second C-position is identified by spreading. "Lexical" is meant in a strict sense, so the spreading-from-left-to-right account may not carry over to geminates in Semitic languages, where geminates arise by derivational processes (the C–position of an empty syll being identified by the melodic content of its right neighbouring root consonant). This would give association a special theoretical status, whereas if spreading always proceeded from left to right, one could claim that identification is pursued automatically. Fortunately, this issue does not affect our analysis of geminates as a special licensing domain.

The sandwiched empty nucleus of a geminate is licensed by this identification mechanism. From the reinterpretation of the Moroccan Arabic data presented in Kaye, Lowenstamm, and Vergnaud (1989/1990), where geminates also constitute a possible domain for proper government, it is clear that geminates must have a special status in the theory. In the 2nd (causative) binyan, the second consonant of a root is a geminate. The following nucleus is licensed to remain uninterpreted by proper government, but not the nucleus before the geminate (e.g., kɨtbu – kɨttbu, but ktɨb – kɨttɨb). In our analysis, geminates receive special status only insofar as they constitute a special case of C–licensing, whereas in earlier accounts geminates had to be attributed to a special kind of governing domain.

4. Assimilation and harmony processes

4.1. The principles of assimilation: spreading, not de-linking

In this approach, direct interaction between positions which are not strictly local (i.e., neighbours on the skeleton) is not permitted. At first sight, this seems to go against a wealth of data from assimilation processes between consonants (voicing, place of articulation). Such effects must be achieved across an empty nucleus — but there are no direct onset-to-onset assimilations across a filled nucleus.[18]

Since assimilation processes are not universal, their occurrence in any language must be regulated by parameter settings. On the other hand, the types of possible assimilation are limited by the inventory of elements (which is not large). The relevant parameters, we propose, are conditions on government. The general form of such conditions is either “element *m* is licensed in position *i* if its governing position *i*+1 also contains/acquires element *m*” or “position *i* containing element *m* can govern position *i*–1 if position *i*–1 also contains/acquires element *m*”. Of course, such conditions do not actively carry out the assimilation; rather, they stipulate that the unassimilated realisation of the phonological domain in question is illicit. The activation of assimilation, we believe, has to do with the language-specific status of certain elements. Assimilating elements have some property which allows them to become visible in some way to a neighbouring syll. Thus in languages with voicing assimilation processes, the tone elements (H,L) in onsets have this kind of visibility; in languages with place assimilation of consonants, the elements (I,U,R) have it. In vowel harmony processes, typically some or all of the elements (I,U,F,$\underline{\text{F}}$) are visible in this way.

Clearly, further details need to be worked out, and we will simply assume for the purposes of the present paper that this approach is basically correct. Given the two (parametric) bans on unassimilated sequences and an activation mechanism, it should be easy to visualise how the vast majority of classic assimilation processes are treated within this framework. In the next subsection we examine an apparent problem with place assimilation.

4.2. Place-assimilation of nasals: an apparent problem

In many languages, nasals assimilate quite eagerly when they are in the immediate vicinity of stops. This poses an apparent problem for our conception of place, especially for the representation of velars, which lack a specific element specifying 'velar'. The only mechanism by which assimilation can be effected is by the spreading of an element from one segment's melodic expression to some other segment. De-linking is not a mechanism of phonology. Hence, a true coronal consonant can never be assimilated to a velar.

Consider the apparent counterexamples of the one Germanic and two Latinate English prefixes *con-* and *in-*, given in (10).[19]

(10)	a.	u[m]purified $^{?}$[n], *[9]	co[m]pare *[n], *[9]	i[m]part *[n], *[9]
		u[m]bounded $^{?}$[n], *[9]	co[m]bine *[n], *[9]	i[m]bibe *[n], *[9]
		u[n]tenable *[m], *[9]	co[n]tain *[m], *[9]	i[n]tend *[m], *[9]
		u[n]doubted *[m], *[9]	co[n]done *[m], *[9]	i[n]dent *[m], *[9]
	b.	u[9]confirmed *[m], $^{?}$[n]	co[9]cur *[m], *[n]	i[9]cur *[m], $^{?}$[n]
		u[9]grateful *[m], $^{?}$[n]	co[9]gratulate *[m], *[n]	i[9]gress *[m], $^{?}$[n]
	c.	u[n]attractive *[m], *[9]	co[∅]erce *[m], *[n], *[9]	i[n]itial *[m], *[9]

Traditional accounts of these assimilations assume that words of the type *unattractive* and *initial* show the underlying form of the prefix because the stem to which they are attached begins with a vowel. However, forms like *coerce* ought to tell us that something else is going on here, especially in view of the phonotactic absence in English of word-initial [V9V] sequences.

We therefore propose that the lexical shape of these three prefixes does indeed have a final velar nasal (L). The assimilations in (10a) are then trivial leftward spreadings of all *U*- and *R*-elements of the stem-initial consonants.[20] The velar cases in (10b) are pseudo-assimilation, i.e., the lack of any assimilating element in both the source and the target segment. What remain to be explained are 1. the sporadic [n]'s in the second lines of the (10a) and 2. the (10c) forms. The [n]'s in (10a) could be due either to "reading aloud" hypercorrect pronunciations (since all these nasals are spelled with an ‹n›), or to a strengthening of the morpheme boundary to a domain boundary, in which case a single final [9] is illicit, and must be supplied with an ambient *R*-element (producing [n]).

The forms in (10c) are particularly interesting, since they manifest two different strategies for getting rid of an illicit single [9]: in the first and third column (*unattractive* and *initial*) the ambient *R*-element must be added to the nasal, while in the second column (*coerce*) the nasal is lost completely. Why two different strategies? Because the sequence [əwɜː] is permissible in English, while the hiatuses [ʌə] and [ɪɪ] are not. Note that the traditional view that these prefixes each have a final coronal nasal [n] has no explanation at all for the *co-* variant. Our analysis, on the other hand, is based squarely on the unlicensable status of [9] except in a geminate or “nasal + homorganic stop” sequence.

It should be clear that, although English examples were used here, the same mechanisms apply to all place assimilations in all languages.

5. Consonant clusters

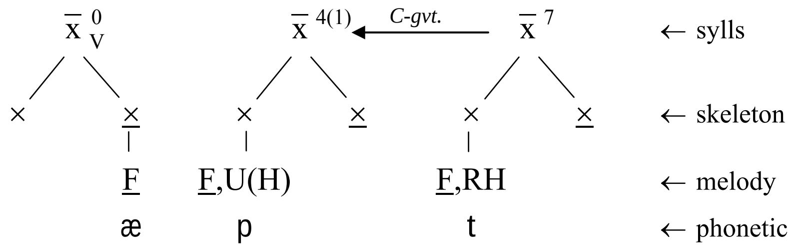
Having outlined the model, we now show how it works with some selected examples which were particularly problematic in previous models of phonology.

5.1. “Onset-to-onset”: the ends of English and German words

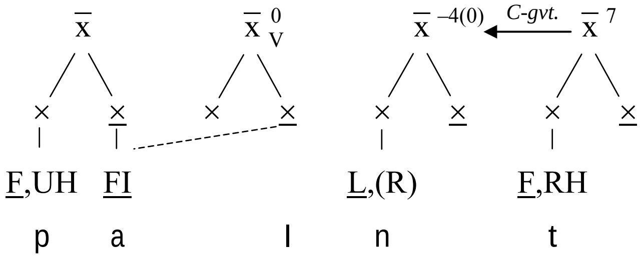
We propose that previous right-to-left onset-to-onset relations involve a parameter setting which allows the onset melody of a particular syll to be visible for the evaluation of licensing in the preceding syll (C-government – see (8b)).

Let us consider consonant sequences at the ends of words in English and German. The final nucleus is licensed by the final empty nucleus parameter; but several of the preceding sylls also contain empty nuclei, all of which are licensed by C-government. In (11) we give some classical complex coda examples:

(11) a. English *apt*



b. English *pint*



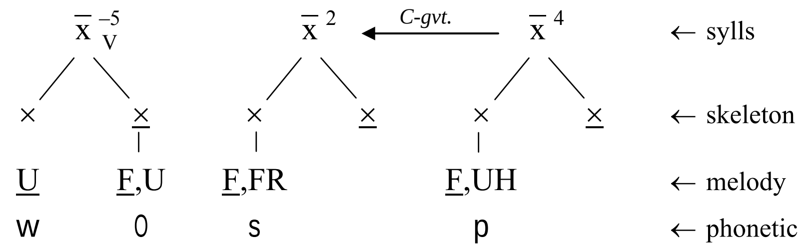
In both cases, the final syll is licensed because it contains a final empty nucleus, but the preceding syll posed a problem to classic Government Phonology. Under the present approach, C-government ensures that the empty nuclei are licensed in a principled manner, and of course prohibits the reverse order of consonants.

Now let us look at sequences of consonants where one is a fricative. When the fricative is followed by a fortis stop, the same analysis as in (11) applies. However, if /s/ were represented as (F,FRH), it would have a melodic strength value of 5, fortis /k/, /p/, /t/ have the values 3, 4 and 7, respectively. So at first sight it would seem that only [st] is possible. However, since voicing assimilation is all-pervasive in English consonant sequences, we can assume a simpler lexical melody (F,FR) with a strength value of 2 for [s] in these contexts. This now permits [sp] and [sk]. Let us formulate this more precisely:

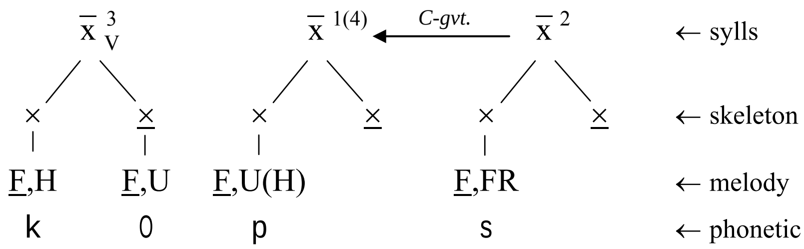
(12) Governed source element hypothesis:
In a C-governing relation, source elements (*H* and *L*) count for strength only if the syll is ungoverned.

This hypothesis formalizes the effects of voicing assimilation and accounts for the apparent (i.e. phonetic, but not phonological) reversibility of final [sp]/[ps] and [sk]/[ks] as in *wasp/copse, risk/fox*; in each case, the second consonant ‘contains’ *H* but the first does not. The representations of *wasp*/*copse* are given in (13) below.

(13) a. English *wasp*

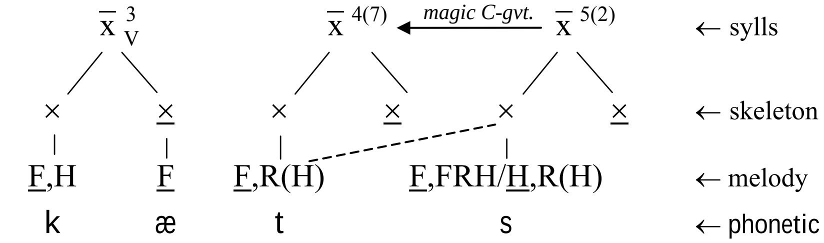


b. English *copse*



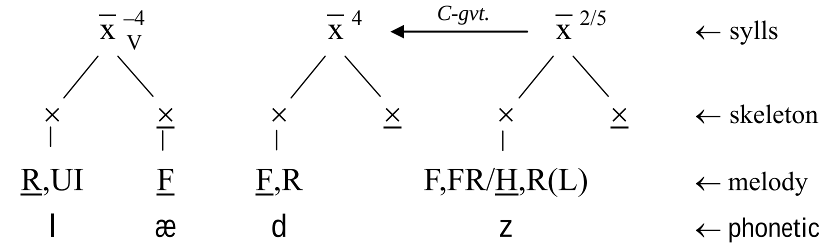
The same holds for final [sk] and [ks]. Let us consider the more problematic cases: [ts] and [dz]. These sequences do not exist in English monomorphemic words, but they abound in plurals. First, consider the representation in (14). Here, we have to assume that the source element *H* which originates from the /t/ is assimilated and therefore also interpreted in the last syll.

(14) English *cats*



Even more problematic is the case of [dz], since there is apparently no source element available which enables us to represent both [zd] and [dz]. There are two strategies. One would be to assume that in these contexts /s/ in fact forms a phonological domain on its own, and the first syll of the cluster is licensed as a (subdomain-)final empty nucleus, the other, which we will explore only briefly, that /s/ is in fact (H,R) and voiced /z/ has an *L* operator.

(15) English *lads*

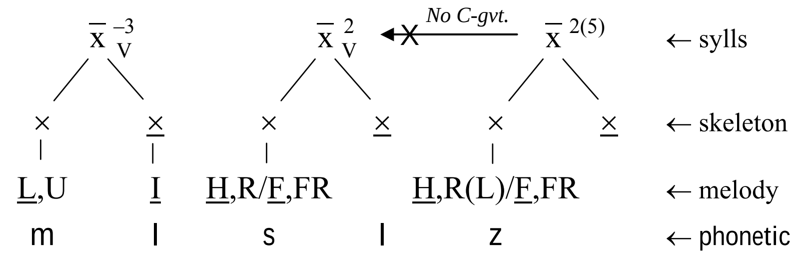


If we have to assume a special licensing mechanism which is not available except across a morpheme boundary, we lose the generality of our account, however, the other voicing-assimilation facts fit in perfectly. On the other hand, if we assume /s/ to involve an *H*-head we capture the fact that /s/ is more common than /z/ in onsets. Voicing takes place only after vowels or if /s/ is part of a post-vocalic consonant cluster devoid of any (conflicting) *H*-operator. Voicing amounts to adding an *L*-operator to the melodic representation of /s/ (H,R), and has exactly the same strengh-boosting effect as an *H*-

operator, being subject to the Governed Source Element Hypothesis. Both accounts have their advantages and disadvantages.

Despite the wide range of final consonant sequences that can be licensed, there remain two classic cases where C-government (correctly) fails: in the [ɪz] variant of the *-s* suffix, which appears after stem-final [s,z,S,Z,(tS,dZ)], and in the [ɪd] variant of the *-ed* suffix after stem-final [t,d]. The first of these cases is exemplified with the word *misses* in (16). Within the present theory there is no mechanism that could license the penultimate syll of *misses*, hence its nucleus must be identified with the default vowel.

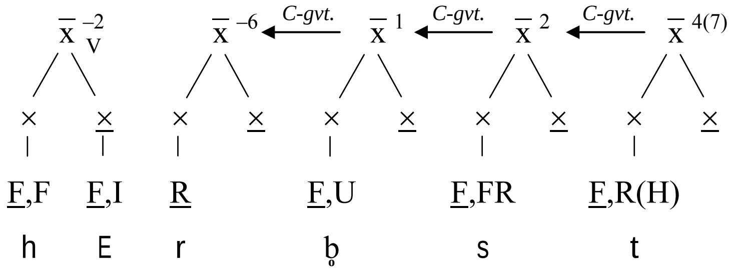
(16) English *misses*



For reasons of space, parallel representations with other stem-final consonants and with the *-d* suffix cannot be given, but they clearly work in exactly the same way.[21]

Finally let us consider the longest word/stem-final sequence of consonants that we have found so far — that in Standard German *Herbst* 'autumn'. Its representation is given in (17).[22]

(17) Standard German *Herbst* ‘autumn’ (cf. English *harvest*)



Each syll of the complex coda (except *r*) can C-govern its left neighbour because each nucleus x̲ is empty and each syll is melodically stronger than the its left neighbour.

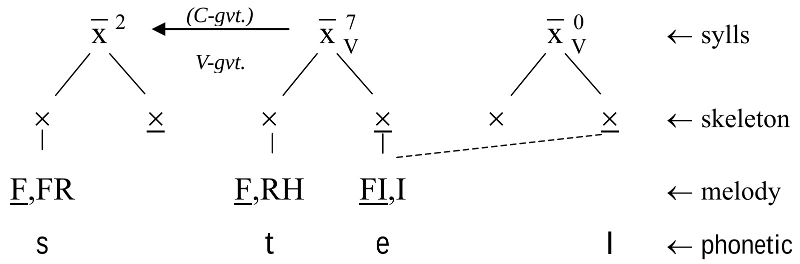
Unsurprisingly, coronal obstruents, the strongest consonants will tend to occupy the last position in a string of sylls with empty nuclei, because the *R*-element provides extra strength. This is probably the reason for the frequency of coronals as suffixes in the languages of the world, and the misconception among linguists that this is the “default” place of articulation.

5.2. *Word-initial* sC *sequences*

There are clearly two classes of so-called word-initial *sC*-sequences; one where *C* is “weaker” than *s* and the other where it is stronger. We propose that the former class in fact comprises contour segments of the usual kind: thus for example s^w in English *switch* is simply a labiovelarised *s* (i.e., *s* with a lazy *U*-head), akin to the other labiovelarised consonants like t^w in *twitch,* k^w in *quick*, T^w in *thwart*. Similar accounts are available for s^l (lazy *R*-head and *UI*-operators) and s^j (lazy *I*-head). The onset S^r as in *shriek* seems to be “what would have been” s^r.[23]

The second class of *sC*-sequences involves two distinct sylls. This is shown in (18) for initial /st/ in the English word *stay*.

(18) English *stay*



The first syll in (18) is V-governed. C-government is not necessary to license the first empty nucleus.[24] What is crucial for the first position of these clusters is that the weaker consonants are excluded; only the quite strong consonant /s/ will be found at the very beginning of the phonological domain. We capture this with the First Syll Requirement in (9), which demands that the very first syll in a domain-initial cluster have a C-governing potential stronger than 0.

The motivation for the First Syll Requirement might be the empty initial CV site of Lowenstamm (1999), which has a melodic value of 0 by virtue of its emptiness. The diachronic filling of the head × of this word-initial empty CV with the melody /e/ in Spanish and French results from a weakening of the C-government capacity of voiceless stops in those languages (probably through the loss of an *H*-operator).[25] As a consequence, the first empty syll could no longer be licensed, and the nucleus of that syll had to be realised phonetically like any other unlicensed vocalic position.

In some languages (e.g., Polish, Greek) the First Syll Requirement is not active, which means that a greater variety of initial consonant sequences is possible. We hope that an analysis of the Polish facts based on right-to-left government of the type outlined above will be possible. The Greek facts, on the other hand, fall out quite easily.

Note that in all these cases, right-to-left government, which in most word-initial cases exactly contradicts the Sonority Sequencing Generalisation, makes exactly the right predictions. In addition, it accounts for *#pt*, *#xt*, etc., but **#tp*, **#tx* in Greek and Serbo-Croatian by the higher strength of the *R*-element, while the Sonority

Sequencing Generalisation has nothing to say about #*pt* and gets #*xt* wrong.

6. Prosodic structure – the foot and its head[26]

Thus far we have not said anything about higher structuring. Sylls as projections of nuclei were claimed to be the mediating nodes for prosodic structure. We cannot lay out a fully specified model of prosody here, let us propose just the minimal setup and leave the answer to upcoming questions for future research.

There is quite surprising agreement among linguists as to the label for higher structural units, i.e., *feet*, considering that there is almost no agreement on what a foot actually is. How much structure can it comprise and what are its exact properties? We will not tackle this task directly, but we will try to advance one small step by considering the consequences of the following principle (which simply follows from the principles of phrase structure):

(19) Foot headedness
Every foot F has a head F°, which also counts as its prosodic head.

This is how sylls are organised into higher structures. The head syll of the foot forms one constituent with its right neighbouring syll; all the other sylls within the phonological domain will be attached to the maximal prosodic projection, the foot itself. Is it relevant to discern domains on a higher level? We believe that it is. Consider the following hypothesis:

(20) Domain opacity hypothesis

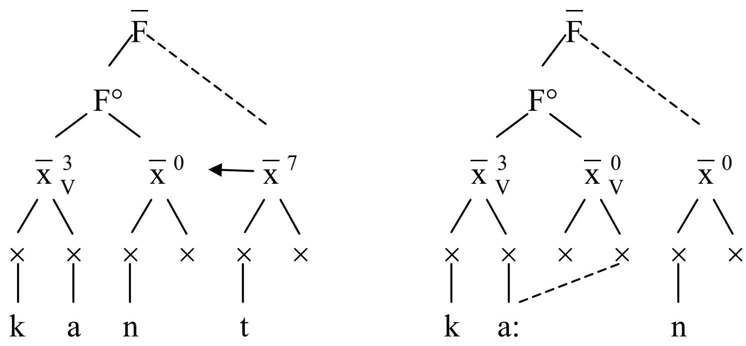
1. Licensing is bounded to a phonological domain; V-licensing (proper government and final empty nuclei) cannot apply to a phonological subdomain.
2. The head of a foot forms a phonological subdomain.
3. Therefore, the syll(s) which constitute the head of a foot cannot be licensed by V-licensing.

This brings in an asymmetry between V and C-licensing. V-licensing (former proper government and the special final empty nucleus-licensing) are bound to domains outside the head of the foot, while C-licensing is carried out just on the level of sylls under adjacency, regardless of higher structure (but still within one phonological domain). This fits in neatly with what we have claimed about the differences in the structural properties of consonants and vowels: consonants never see each other; only their melodic values percolate up to their dominating node, the syll, where they may be evaluated against each other. On the other hand, vowels project to sylls and relations between vowels are calculated via structure per se. So it is not really surprising that V-licensing is sensitive to substructure in the prosodic domain, while C-licensing is not.

Let us examine some consequences of these claims. First, it falls out naturally that every phonological domain must have at least one phonetically interpreted nucleus. Every foot must have a head and the nuclei of that head cannot be V-licensed, hence they have to be realised phonetically. They could, in principle, be C-licensed; in a moment we will see how this is prohibited, at least for German.

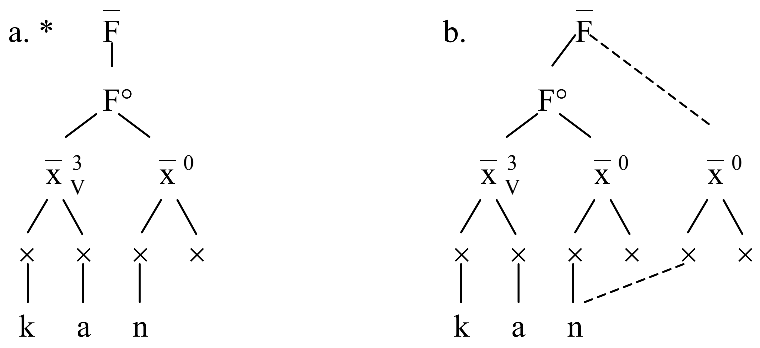
What we still need is an exact definition of what the head of a foot looks like, or more specifically, how many sylls may or must be in the head of a foot. This may be subject to parametric variation, for a language like German or Italian, where we clearly face long vowels and geminates in complementary distribution (at least phonologically), we assume that there must always be exactly two sylls forming the head of a foot. For illustration, consider the following examples from German. (The dashed association lines from melodic elements to the equivalent positions in the right-neighbouring syll indicate vowel length/diphthongs and gemination, dashed association lines from the F-bar node to sylls outside the head of the foot indicate that those sylls do not have an interpreted vowel, and thus may not count for metrical purposes.)

(21) a. German *Kant*[27] 'Kant' b. German *Kahn* 'barge'



In (21a) the Domain opacity hypothesis is fulfilled by virtue of C-government. The second syll contains an uninterpreted vowel which is not licensed (and cannot be) by V-government or final empty nucleus-licensing. The first syll has an interpreted vowel, the short /a/. (21b) differs from that in the length of the vowel: the /a/ is long and spreads onto the second syll rendering both V-positions as interpreted. The third syll can always be licensed as a final empty nucleus. That far we considered mono-syllabic words (feet) of two types: one with a complex coda (CVCC) and one with a long vowel (CV:C). What about other types? A short open syllable is never attested to be a licit word/foot in German, a long open syllable would be the minimal constellation, two sylls with a realised vowel (or diphthong, the representation would be the same). A closed short syllable with two sylls should not exist according to the Domain opacity hypothesis since the uninterpreted nucleus would have to be licensed as a final empty nucleus, which counts as V-licensing (impossible under that configuration). Consider an example like *kann* 'can' in German, in (22a) as two-syllablic and in (22b) as a configuration with 3 sylls:

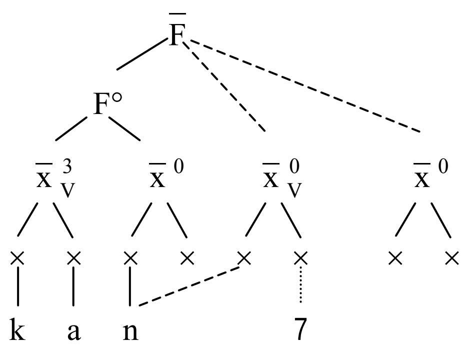
(22) German *kann*
‘can’



The first representation is illicit since the vowel of the second syll should either be licensed or receive a phonetic interpretation. Hence, the only way how to represent *kann* or any other mono-syllabic word with CVC is with a geminate. In (22b) the second syll receives its license by C-government (the special geminate case), which is an available option.

Consider now a richer morphological structure, as for example *Kanne* in German which minimally differs from (22) in that there is a schwa at the end of the word. It is is in principle a nominal stem, however it can be assumed that the schwa indicates a historical derivation, indicating one of the feminine inflectional classes.

(23) German *Kanne* ‘jug, pot’



For the schwa to show up at the end of a word we assume that there is an additional syll present, which is totally devoid of melodic specification. (Maybe it should be considered as an inflectional affix without melodic content.)[28] This syll only has effect on licensing properties of the structure, its vowel position is licensed by a final empty nucleus, but this also means that the vowel of the third syll is unlicensed, hence realised as schwa. Licensing of the other position does not differ from the structure we presented in (22).

This analysis also sheds some light on geminates in general. We assume that there may be several sources for geminates, some of them purely morphological (as is generally assumed for Semitic or Finno-Ugric languages), some of them as a mere reflex of the interaction between licensing and higher-level structure.

If this analysis is on the right track, it will also have some interesting consequences to a question which is rarely addressed: Why do languages *sound* so different? In other words, where is the structural difference between languages which regulates their different prosodic output? Under this model it will be the result of a modular interaction of independent parameterised principles, determining the licensing configurations of empty vowel-positions and the shape heads of phonological domains will take.

Conclusion

The specific model of Government Phonology that we have outlined in this paper still leaves a number of questions open for further investigation. In particular (as we have been at pains to point out), several of the mechanisms that we have postulated are still solely internally motivated, even though they achieve observational adequacy.

What we hope to have shown is that (a) the basic CV structure is well founded and can be integrated into a strict formal generative model in the form of the syll constituent, and (b) government operates exclusively from right to left. With a slight revision of the system of elements and the (still observational) definition of individual strength properties of these elements, we can capture a wide array of phenomena relating to consonant clusters which remained mysteries during the past years of progress. This required a redefinition of government as the basic relation of phonological structure, in particular government operating through sylls only, and with two manifestations according to the syll's two sub-constituents, consonants and vowels.

Proliferating government to the edges of phonological domains provides good means to explain the special behaviour of the first onset and the final coda of a domain without invoking any additional mechanics. Finally, exploiting the concept of sylls and domains, we have taken a first step towards the integration of phonological structure into higher prosodic domains.

Notes

* We are grateful to many people for inspiration and/or comments: Mathias Böhlinger, Monik Charette, John Harris, Christian Huber, Harry van der Hulst, Jonathan Kaye, Jean Lowenstamm, Ádám Nádasdy, Markus Pöchtrager, Stefan Ploch, Péter Rebrus, Elli Rieder, Grażyna Rowicka, Tobias Scheer, Péter Siptár, Péter Szigetvári, Miklós Törkenczy and many participants at various conferences over the past three years. By the time this is printed, neither they nor we will agree with everything (perhaps even anything) written here. Nevertheless we assume full responsibility for this product of our misspent evenings, and absolve one and all from any part in the blame.

1 Cf. Harris 1990, [1992](1997), 1994, Kaye, Lowenstamm and Vergnaud 1989, Ploch 1999, Scheer 1996.

2. Only right-headed structures support abductive hypothesis-based parsing, which we assume is a basic ingredient, alongside redundancy, of the robustness of phonology.

3. Our choice of the term *syll* was motivated by the fact that this constituent is in traditional terms a CV-syllable. Since the maximal syllable constituent is now shorter than before, its name has been shortened too. The arguments of Kaye, Lowenstamm, and Vergnaud (1985, 1989, 1990) against the syllable as a constituent no longer apply here; once traditional "branching onsets", "branching nuclei" and post-nuclear consonants ("codas") are abolished, constituents are right-headed, and "heavy syllables" comprise more than one syll.

4. We are aware that the reverse order, i.e., VC, as proposed by Szigetvári (1999), could be equally possible. However, there are two reasons for us to believe that

the proposed order CV is the adequate one. On conceptual grounds there is the (possibly disputable) analogy of right-headedness. Methodologically we show that under the proposed assumptions a model can be gained which explains quite an extended array of data. However, this model of licensing and government can only be maintained under the syll = CV hypothesis. We know that these arguments are only system-specific and therefore inappropriate to guide any decision between CV and VC. Hopefully, future work will shed more light on this dispute.

5. This does not equate to the claim that CV is the minimal pronounceable *utterance* of every language; many languages have minimal word requirements, cf. Szigetvári (2001).
6. We suspect that there will be qualitative differences between lower and higher-level domains, especially those related to prosody. But for present purposes, at the word level, domains can be considered to be of the same type.
7. The tensing of long vowels and the spirantisation of intervocalic stops can now be seen as parallel processes: functional sandwiching effects. An empty onset sandwiched between the two vowel positions (which in our terms constitute a long vowel) acts as an *F* operator in the neighbouring nuclei (producing, for example, the tensing of long vowels in German), and two full nuclei sandwiching a stop consonant may add an *F* operator to the consonant's melodic representation (giving intervocalic spirantisation).
8. For example, in German the floating *I*-element which induces umlaut (in alternating stems) is introduced by a suffix, but because it must be realised in a nucleus already containing melody it may finish up being attached several sylls to the left, e.g., in *Papst* 'Pope', plural *Päpste*, where it jumps over three sylls. That this cannot be analysed as purely lexical can be seen from the fact that umlaut is still productive in certain contexts ("Hallo" – "Hallöchen").
9. This occurs, for example, in Mòoré vowel capture (umlaut), where the quality (*I*, *U* or *A*) of a suffix vowel is "captured" into the stem vowel (with contextual restrictions) and is not realised in its original suffixal position except phrase-finally, when it is realised in both positions. In cases where the stem vowel is originally a short monophthong, the resulting short diphthong is a contour segment.
10. This term seems to be fully equivalent to what other phonologists have defined as a "phonological expression". Cf. Kaye 1992, 1997; Ploch 1995, 1999a, c.
11. Note that whenever *F* has the phonetic interpretation of the traditional *A*-element, it must be the head; as an operator it functions like the previous ATR-element. However, whether *U* or *I* are heads or operators does not really matter, they will take head status whenever no other element is head. In a melodic expression consisting of both *I* and *U* and nothing else, the vowel with a *U* head will be more rounded than the *I*-headed vowel, as in Norwegian outrounded ($\underline{\text{U}}$,I) *versus* inrounded ($\underline{\text{I}}$,U).

12. While the lack of *R* in (lexical) nuclei is still only an observation, it is clear that languages exploit this. In cases where an unexpected consonant is inserted (by whatever mechanism), such as English and Austrian German "intrusive *r*", it is often a coronal because coronality is uniquely linked to an onset position. See §4.2 below for a similar effect in nasal consonants.
    We are aware that this contradicts other proposals which merge the R-element with the vocalic A-element. The architecture of elements we are proposing here speaks against this. R is the most significant but not the most prototypical 'consonantal' element. Therefore we assume the functional element to represent 'A'-ness in vowels and 'stops' in consonants, whereas the R element is a good parsing trigger that a position is definitely consonantal. The empirical evidence for the R=A account (cf. references cited above, especially Ploch 1993) taken from Schwabian dialects (but the phenomenon extends to almost all Southern German variants) can be undermined on phonological grounds. Fact is that non-low vowels before /R/ are lowered or form falling diphthongs with an a-schwa, which apparently calls for a spreading analysis. However, an a-schwa can never be stressed and they also never undergo processes like 'Umlaut'. In our opinion, the most straightforward way to explain this is that there never is an A/F-element present and the lowered schwa-sound is a result of the phonetic interpretation of the bare R-head in the adjacent consonant position to the right, a phenomenon which is also found with 'spreading' of nasality from nasal consonants to neighboring vowels.
13. At this stage, we have no concrete new proposal for replacing the traditional melodic tiers. However, it seems that they are likely candidates for revision in the future development of the present approach. In particular, since the relatively free "spreading" of elements to skeletal points is no longer possible (but must be mediated by sylls), there seems to be little call for the complex structure of melodic tiers assumed by most researchers (including ourselves) up to now.
14. Austrian German, for example, has no voicing contrast. Single obstruents are partially voiced, geminates voiceless. English has "H-(de)voicing" and French has "L-voicing" (cf. Harris 1994). The initial voiceless, lightly aspirated [k^h] is a contour segment which parallels [p^f] and [t^s], but which is probably synchronically reanalysed as having a lazy H element.
15. We assume this to be the case for the Odawa dialect of Ojibwa (Kaye 1975 based on Piggott 1974).
16. The phonetic realisation of a melodically empty position is most often [ʔ] in an onset and [ə] in a nucleus, but variations are possible across languages. Among other things, this means that the empty phonological structure inserted by speakers to fill pauses is (maybe universally) $\overline{\times}(\times(),\underline{\times}())$, which translates into English as [ʔə] (orthographic ‹er›) and Austrian German as [ʔɐ].

17. This reinterpretation of the final empty nucleus parameter is particularly felicitous in cases such as Koromfe, where final empty nuclei are licensed on the word level but not on the phrase level (where, in the last resort, an *I*-element must be "inserted" if neither *I* nor *U* are available from a local source, see Rennison 1997 for details).
18. For assimilations such as Nyangumarda /p/ → /m/ in the tense/aspect marker /lVpV/ (where V is a harmonised vowel from the set /i, u, a/) when the following morpheme begins with a nasal consonant, we must assume that nasality actually spreads to the intervening V. Unfortunately, we have no way to check whether this vowel is phonetically nasal or not. Note also that this nucleus is at least lexically empty (cf. Rennison 1987).
19. We make no claim that all of these words are analysed by present-day speakers as having a prefix; we are only interested in the place assimilation of the nasal.
20. The same goes for words with all other stem-initial consonants, which we have only omitted from (10) for simplicity of presentation.
21. A parallel analysis of Standard German schwa in the penultimate sylls of *arbeitet*, *redet*, etc., is also available. However, that we do not yet know why words like *rechnet* also have such a schwa. In Austrian German the words of the type *arbeit(e)t*, *red(e)t* always have a geminate, and the equivalent of *rechnet*, namely [rEçN̩d̥], is well-formed according to the principles outlined above — i.e., it corresponds to SG **rechent*.
22. In Austrian German the /r/ of this word becomes a low schwa-like offglide [6], i.e., the representation and licensing of the word are the same as in Standard German, but in addition the *R*-element of /r/ spreads leftward into the neighbouring nucleus, forming the contour segment [E6].
23. We have no account, as yet, for why its *s* is realised here as [S]. Note the parallel with German initial /S/+C, which were historically /s/+C.
24. If it were, it would provide a natural explanation for why in English and German only fortis stops can form *sC*-clusters at the beginning of a domain. However, we do not yet know how to motivate such a move. That V-government is a necessary prerequisite can be shown by data from Italian: there is no /s/-voicing in Italian, but there are *sC*- and *sC*-liquid clusters with both fortis and lenis stops. One of these classes of stops would not exist in that peculiar configuration if these clusters in Italian were not licensed by V-government alone.
25. It seems that Latin had H-(de)voicing, and we assume that present-day French and Spanish have L-voicing.
26. The ideas presented in this section are basically a reformulation and further implementation of considerations Jean Lowenstam presented in his classes in Vienna and a combination of them with John Harris' "Foot as an informational domain." We are grateful to both of them, any misconceptions here would stem from misinterpretations on our side.

27. *Kant* is a proper name in German, rather than a genuine word. However, it may be part of a complex verb form, e.g., the participle *be-kannt* ‘known’. The geminate in orthography we regard as not being present in the phonological representation. A similar word, *Kante* ‘edge’, involves more structure, in particular a melodically empty syll, which consumes all the benefits of final empty nucleus-licensing and forces a schwa to come up after the /t/, as the interpretation of an unlicensed empty vowel-position.
28. Notice that if we had an affix like *–en*, *-et* or *-er*, these would structurally not differ at all. The only difference would be an *L*-head, an *F*-head (with *R*) or an *R*-head in the C-position of the relevant syll. Final empty nucleus-licensing operates as usual, and the schwa will then show up anyway, unless a C-licensing configuration allows for “schwa-deletion”, or better licensing of a vocalic position.