UNIVERSITÉ DU QUÉBEC À MONTRÉAL

THE PHONOLOGY OF EMPTY CATEGORIES

THÈSE

PRÉSENTÉE

COMME EXIGENCE PARTIELLE

DU DOCTORAT EN LINGUISTIQUE

PAR

PAUL JOHN

AVRIL 2014

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

LA PHONOLOGIE DES CATÉGORIES VIDES

THÈSE

PRÉSENTÉE

COMME EXIGENCE PARTIELLE

DU DOCTORAT EN LINGUISTIQUE

PAR

PAUL JOHN

AVRIL 2014

ACKNOWLEDGEMENTS

My greatest debt of gratitude goes to my thesis supervisor, Mohamed Guerssel, who did a stellar job training me in Government Phonology and guiding me in developing the ideas contained in this thesis. Mohamed challenged me to be more rigorous when my argumentation was faulty or my examples flimsy. By pushing me and by providing key suggestions, Mohamed has ensured a considerably improved final analysis. Frankly, I cannot overstate how appreciated his contribution has been. Though not reflected in the content of the thesis, I appreciate likewise his advice on some outstanding jazz recordings that I might not have discovered without his insight. Mohamed, I think I can honestly say you have always given me sound advice.

At various stages in the development of this research project, John Lumsden and Renée Lambert gave invaluable feedback and posed perceptive questions that helped me to improve my thinking. I thank you both enormously for your time and patience. Feedback from fellow graduate students in the context of the *séminaire de synthèse* and the *atelier de recherche* was also much appreciated. At the final stage, I was inordinately lucky to enlist Heather Goad and Heather Newell, complementing Mohamed Guerssel and John Lumsden as members of my final committee of readers. I am certainly grateful for Heather Goad's willingness to somehow fit me into an already insanely busy schedule. In a less official role, I also benefited from probing discussions with Walcir Cardoso, especially as to the eventual applicability of my ideas to second language phonology.

I alone am to blame for any residual flaws or infelicities.

For my much-missed father,
Brian John,
and my much-loved wife,
Ying Su

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

RÉSUMÉ

Les catégories syllabiques vides (attaques, noyaux et codas) ont de par le passé été représentées comme étant reliées à des points de squelette auxquels aucun contenu mélodique n'est associé. Les attaques vides constituent une exception à ce genre de représentation. Dans le cadre de la phonologie du gouvernement (ainsi que dans d'autres cadres théoriques) deux types de représentation ont été proposés pour les attaques dites vides : avec ou sans point de squelette. Cette distinction est motivée par le comportement variable des attaques vides en termes de certains processus morpho-phonologiques, notamment à l'initiale de mot, comme par exemple en français où certaines formes commencent soit par une voyelle soit par un h-aspiré. Cependant, le présent travail établit qu'une distinction représentationnelle doit s'appliquer aussi bien aux noyaux et codas qu'aux attaques vides. De plus, une distinction basée sur la présence versus l'absence d'un noeud de racine est plus apte à capturer les différences de comportement observées.

Les différences de comportement observées comprennent : i) la capacité d'une catégorie vide à rester vide; et ii) la capacité d'une catégorie vide à être la cible d'un processus de propagation locale, soit d'éléments mélodiques individuels, soit d'un segment voisin au complet.

Le choix du nœud de racine comme déclencheur de la variation n'est pas gratuit. La raison d'être du nœud de racine dans une représentation est de regrouper tous les éléments mélodiques faisant partie d'un segment et de relier l'ensemble au palier des points de squelette. Lorsqu'un nœud de racine non spécifié est présent dans une représentation, nous démontrons que le nœud de racine peut soit déclencher l'insertion de mélodie par défaut (par exemple, d'un cheva ou d'un coup de glotte), soit d'attirer des éléments individuels de mélodie d'une source voisine. Sans le nœud de racine, par contre, la catégorie vide ne contiendrait pas l'ingrédient essentiel pour le déclenchement d'une insertion par défaut, ou pour la propagation d'éléments individuels. Néanmoins, une alternative demeure possible : la présence d'un point de squelette permet à la catégorie vide d'être la cible d'une propagation du noeud de racine figurant sous une position voisine. Dépendamment du degré de dégénération d'une représentation donnée, nous réussissons à caractériser adéquatement le comportement des catégories vides en général.

MOTS-CLÉS : catégories vides (attaque, noyau, coda), insertion par défaut, propagation locale, Phonologie du gouvernement, Théorie des éléments

ABSTRACT

Empty syllabic categories (onsets, nuclei and codas) have generally been represented as constituents containing timing slots and by definition no components of melody. The exception concerns empty onsets, which in Government Phonology (and other frameworks) have been posited as having two possible representations: either with or without a timing slot. The distinction is motivated by the variable behaviour of empty onsets with regard to certain morphophonological processes, notably at the beginning of vowel-initial or *h-aspiré* forms in French. This thesis demonstrates, however, first that a representational distinction must be extended to empty nuclei and empty codas, and second that a distinction based on presence versus absence of a root node better captures the observed differences in behaviour.

The kinds of differences in behaviour that are found concern: i) whether the empty category can remain empty or not; and ii) whether the empty category can potentially be a target of local spreading of individual melodic elements from neighbouring segments or else only of the neighbouring segment in its entirety.

The choice of the root node as locus of variation is not arbitrary. The root node is a docking site that groups together the disparate melodic elements of subsegmental structure and links this melody to the timing tier. In the absence of specified melody, it is my contention that the root node triggers default insertion (e.g., in the form of schwa or glottal stop) or else it attracts melody from a neighbouring segment via double association of individual elements. In the absence of a root node, the empty category no longer contains the requisite prosodic unit to generate default melody, nor can it harbour select elements from adjacent segments. One option remains open to it, however: the presence of a timing slot allows the empty category to be a target of spreading via double association of the root node under an adjacent position. The proposal to represent empty categories with two degrees of degeneracy based on the presence versus absence of a root node thus proves to be an effective predictor of the type of process the empty category can undergo.

KEYWORDS: empty categories (onset, nucleus, coda), default insertion, local spreading, Government Phonology, Element Theory

INTRODUCTION

Two issues have dominated the phonological landscape since the early nineties. First, work within the framework of Optimality Theory (Prince & Smolensky 1993, McCarthy & Prince 1995) has focused on the *grammatical system* that determines surface forms rather than on *phonological representation*. In brief, OT proposes that surface forms are selected based on how they perform with regard to a set of universal violable constraints. These constraints are organized into language-specific ranking hierarchies such that, for identical input, different languages may select different optimal output. This system of determining surface forms is a departure from earlier rule-based approaches (Chomsky & Halle 1968) or approaches based on principles and parameters (e.g., Government Phonology: Kaye, Lowenstamm & Vergnaud 1990). Although constraints necessarily make reference to aspects of phonological representation (see ONSET, NOCODA, FAITHVOICE, and so on), OT does not propose any innovations regarding representation. Indeed, since input and output forms in an OT tableau are conventionally presented as strings of IPA phonemes, the unstated assumption is that subsegmental and prosodic structure is immediately apparent and uncontroversial.

Second, much work in the past two decades has focused on the motivations behind phonological phenomena. In particular, work in Laboratory Phonology (see the *Papers in Laboratory Phonology* series, starting with Kingston & Beckman 1990) and related frameworks (e.g., Grounded Phonology: Archangeli & Pulleyblank 1994; Functional Phonology: Boersma 1998; or Hayes, Kirchner & Steriade 2004) has re-asserted the functional in phonology, as previously championed by Natural and Natural Generative Phonology (Stampe 1969, 1979; Vennemann 1974; Hooper 1976).

Specifically, these approaches attribute phonological phenomena to the dual (and often conflicting) pressures of articulatory ease and perceptual salience. The general aim is to replace abstract cognitive explanations of phonological processes with concrete substantive ones.

In exploring empty syllabic categories in phonology, this thesis goes against both aspects of the prevailing current in phonology. First, it is concerned primarily with the nature of phonological representations rather than with the grammatical system. A preoccupation in phonology with the nature of prosodic and subsegmental representation characterized the decades immediately following the publication of SPE (Chomsky & Halle 1968), so in a sense my work harkens back to an earlier period. In the wake of SPE, phonologists mounted arguments in favour of the syllable and of syllabic constituents such as onset/rhyme/nucleus/coda (Fudge 1969, 1976; Hooper 1972; Vennemann 1972b; Kahn 1976; Selkirk 1982), as well as in favour of higher levels of prosodic structure such as the foot (Liberman & Prince 1977; Hayes 1981; Halle & Vergnaud 1987), culminating in the multilayered Prosodic Hierarchy (e.g., Nespor & Vogel 1986). In the other direction, a layer of temporal representation was proposed between the syllabic constituents and the segments themselves, represented as either CV-slots (McCarthy 1979; Halle & Vergnaud 1980; Clements & Keyser 1983) or X-slots (Kaye & Lowenstamm 1984; Levin 1985).

The model of syllable structure I adopt here is the one used in Government Phonology (Kaye, Lowenstamm & Vergnaud 1985, 1990), which constitutes the theoretical framework for my investigation. In this model, two traditional elements of representation are absent: the coda constituent and the syllable itself – see Figure 0.1. Post-nuclear consonants are represented as rhymal adjuncts (although the term 'coda' continues to be used informally). The syllable level of representation is simply left out; instead rhymes are represented as projecting up directly to the foot level, and onsets fail to project further. [1]

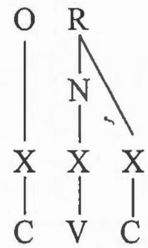


Figure 0.1 *Model of syllable structure*

In the model in Figure 0.1, timing slots are represented by a series of 'X'es; CVC stands in for individual segments, represented in SPE as unstructured feature matrices. Innovative proposals have of course been put forward to revise our understanding of subsegmental structure, notably in the context of Feature Geometry (Clements 1985b; Clements & Hume 1995), which draws on insights into representation originating in autosegmental phonology (Goldsmith 1976). A full-blown feature-geometric representation involves a hierarchically organized tree structure, including a root node at the top dominating further class nodes (the laryngeal, supralaryngeal and place nodes). To these nodes, as shown in Figure 0.2, attach the traditional distinctive features themselves (F_1, F_2, F_3), each a terminal specification on a separate tier.

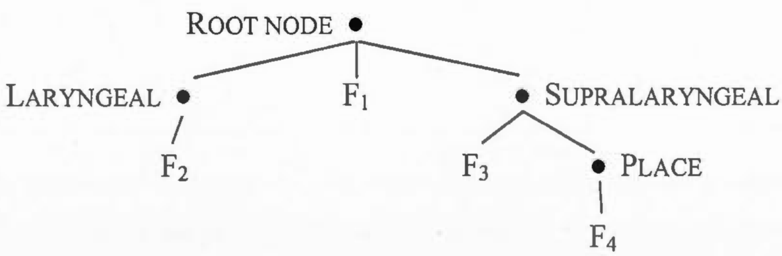


Figure 0.2 *Feature-geometrical representation*

[1] Note also that Government Phonology insists syllable structure is included in lexical representations, since it is not necessarily predictable. In French, for example, liaison takes place in *les oiseaux* [lezwazo] 'the birds' but not in *les ouaouarons* [leØwawarɔ̃] 'bullfrogs'. This fact can be accounted for if the glide [w] is lexically attached in the former case to a nucleus, but in the latter to an onset.

Essentially, in Feature Geometry, the root node serves as a docking site that unifies all subsegmental content and as a link to the timing tier; class nodes serve as docking sites for a portion of segmental content only. As Clements and Hume (1995) demonstrate, the root node facilitates the representation of long segments (a single root node connected to two timing slots) and of short contour segments (two root nodes connected to a single timing slot). It also permits one to represent phenomena such as total assimilation. For example, in Italian, a historical process has replaced sequences of two stops (e.g., [kt]) by a single long stop (e.g., [tt] in *dottore* 'doctor'). With a feature-geometrical representation, this process can be captured quite straightforwardly as delinking of the root node for [k] and regressive spreading of the root node for [t] to the preceding empty timing slot, thus creating a geminate long consonant. Inclusion of class nodes in subsegmental representation, on the other hand, permits instances of partial assimilation, where only a subset of features spreads to an adjacent segment. Likewise, these nodes enable processes of partial deletion (e.g., laryngeal neutralization – Lombardi 1995b) to be captured with a single operation. Finally, the presence of terminal features on their own tiers permits these to spread individually (i.e., single-feature assimilation).

I will be adopting a simpler representation than the one in Figure 0.2. This involves only a root node to which are directly attached melodic primitives (which in Government Phonology are 'elements' rather than features). The question of whether a more complex hierarchical representation is required with intervening class nodes is open to debate.[2] Certainly, what is crucial for my purposes is to insist on the presence of the root node in segmental representations. In particular, the root node will be important as a potential component in the representation of empty syllabic categories.

[2] See Harris (1994) and Harris and Lindsay (1995) for a proposal on how elements might be organized hierarchically under class nodes. However, see also Ploch (2003) for a critique of such highly articulated feature-geometrical representations.

The main argument is that empty categories (i.e., empty nuclei, onsets and codas) can be represented with varying degrees of degeneracy, whether from one language to another or within the same language. Evidence will be presented specifically for two types of empty syllabic categories: i) nuclei, onsets and codas with a timing slot and root node (though no segmental content); and ii) nuclei, onsets and codas with only a timing slot. To demonstrate with reference to nuclei, the two types of empty category are shown in Figure 0.3 below, along with a filled nucleus.

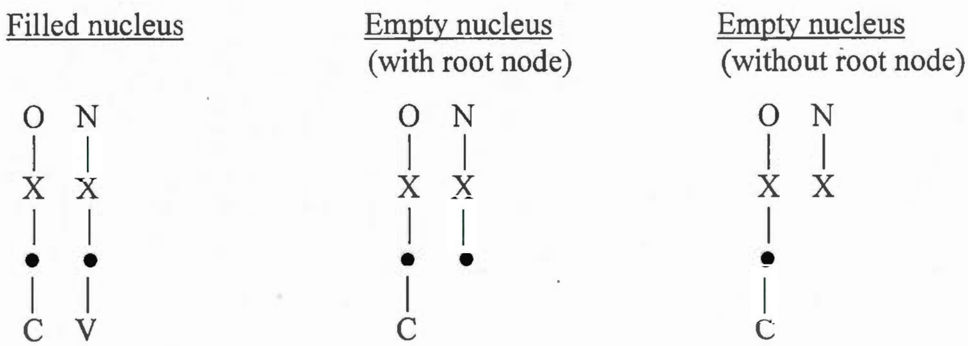


Figure 0.3 *Filled and empty nuclei*

The two types of empty category can be either lexically specified or derived. For example, a filled nucleus can become an empty nucleus either with or without a root node through a lenition process involving delinking of i) only the content of the root node or ii) the root node and associated content together. Similarly, an empty category with a root node might itself undergo a lenition process that delinks its root node. Conversely, an empty category may be targeted by a fortion process that inserts an unspecified root node. Whether a particular surface empty category corresponds directly to a lexical representation or is instead derived is not always easy to establish and must be evaluated on a case-by-case basis.

Regardless of whether lexically specified or derived, the degree of degeneracy of the representation has consequences for the behaviour of the empty category. An empty category with a root node is a target for insertion of default segmental content, such as a schwa or a glottal stop; it is also a potential docking site for a single element spread from an adjacent segment. By virtue of its just containing a timing slot, an empty category without a root node, on the other hand, is a potential docking site only for an adjacent root node. Essentially, then, the representation has consequences for the derivation: presence or absence of a root node determines whether an empty category can receive default melody; it also determines the type of spreading process that can occur. In a nutshell, the burden of explanation shifts from the grammatical system onto the representation itself, recalling McCarthy's (1988) view that "if the representations are right, then the rules will follow."

When similar forms are treated differently in different languages, it is of course possible that the variation stems from the grammatical systems: the languages may simply have different processes, parameter settings or constraint rankings. When apparently identical cases are treated differently within the same language, however, only a representational distinction can account for the difference in behaviour.[3] It is not possible that one process, parameter setting or constraint ranking applies some of the time and another the rest of the time. Crucial support for the representational approach thus comes from distinctions within one and the same language.

Another way in which this thesis runs counter to current tendencies is that it adopts a cognitive/mentalist rather than a functional (articulatory/perceptual) stance. (For a defense of the position adopted here, see Kaye 1989, Hale & Reiss 2000, and van der Hulst 2003; see also Odden 2006 and Blaho 2008 for a radically substance-free view.) Necessarily, any research into empty categories will be cognitive rather than functionalist, since at times the only substance an empty category has is mental substance. Empty categories may lack physical phonetic reality, but they play a role in representations and processes all the same. Certainly, the presence or absence of an

[3] One alternative is to anoint certain forms with lexical markings that set them apart. Tranel (1992), for example, attributes the special behaviour of French *h-aspiré* words to exception features. Being purely stipulative, such *ad hoc* solutions are a last resort that is best avoided.

empty root node in the representation of an empty category is purely abstract/ cognitive. If an empty category behaves differently based on whether it contains a root node, this is hard to attribute to either articulatory ease or perceptual salience.

To the best of my knowledge, empty nuclei were first proposed by Fudge (1976) to account for consonant clusters in Bella Coola and by Anderson (1982) to account for schwa-zero alternations in French. Subsequently, Giegerich (1985), Wiese (1986), and Hall (1989), using various autosegmental paradigms, posited what amounts to an empty nucleus to represent schwa in German. Likewise Spencer (1986) employed the notion of an unspecified V-slot on the CV-tier (i.e., an empty nucleus) to account for yers in Polish. Also in the context of CV phonology, Clements and Keyser (1983) proposed a syllable-initial C dominating zero melody (i.e., an empty onset) to account for the exceptional behaviour of *h-aspiré* words in French. Empty codas were first proposed explicitly as the representation of the initial member of a geminate by Kaye, and Lowenstamm (1984) and by Kaye, Lowenstamm and Vergnaud (1990) in the context of Government Phonology.

More than any other theory, Government Phonology (henceforth GP) has made empty syllabic categories an integral part of its theoretical framework. Consequently GP is adopted as the underlying paradigm for the proposed study, and the thesis can be viewed as building on and modifying the representation of empty categories established in the GP framework. By expanding the range of representations, we are able to account for differences in the behaviour of empty categories in identical contexts. Essentially, these differences in behaviour constitute a central problem to which representational distinctions are the solution. More broadly, what is being proposed is that the capacity to posit varying degrees of degeneracy in empty categories is part of the representational arsenal supplied by Universal Grammar. Not all languages make use of degenerate structures (empty onsets, nuclei and codas), and among those that do, some select only one form of empty category (either with or without a root node), whereas others use both forms. Crucially, the type of empty category employed in a language has implications for how the category behaves, in particular with respect to whether it remains empty or not, and when it does not remain empty, with respect to how it is eventually filled.

The view promoted in this research project is not unprecedented: the notion of different degrees of degeneracy has its precursors in Head-Driven Phonology (van der Hulst & Ritter 1999), specifically as concerns empty onsets and nuclei, and in work by Durand (1995) and by van Oostendorp (1995, 1999), concerning empty nuclei. Nonetheless, this thesis constitutes a particularly thorough-going investigation of empty categories, encompassing degrees of emptiness in onsets, nuclei and codas. Certainly, to my knowledge, this is the first time the issue has been raised of the consequences of the degree of degeneracy for the type of spreading processes that target empty categories. In a critical spirit, the thesis also challenges the GP account of the licensing of empty categories. Briefly, according to GP, empty categories (and particularly nuclei) need to be licensed by an adjacent expressed constituent to remain empty; in my view, empty categories may remain empty if they lack a root node.

In Chapter I, the central tenets of GP are presented in greater detail. Particular emphasis is placed on the GP view of empty categories so as to establish a base for my own proposals. A preliminary account of elements (the privative units of subsegmental structure in GP) is included: special emphasis is placed on the issue of elements that can be inserted by default into an unspecified root node. This issue is expanded on in Chapter II, which is devoted to Element Theory. Subsequent chapters address empty onsets, nuclei and codas in turn, with examples drawn from a range of unrelated languages to demonstrate the recurrent need for a distinction in how empty categories are represented.

CHAPTER I

GENERAL BACKGROUND

1.1 Theoretical framework: Government Phonology

The core texts of Government Phonology are Kaye, Lowenstamm and Vergnaud (1985, 1990), Kaye (1990), and Charette (1990, 1991). In addition, Harris has proposed important refinements of certain aspects of GP (Harris 1994, 1997). In the following review, for ease of exposition distinctions between seminal texts and Harris' extensions will not always be specified.

GP employs a Principles and Parameters framework (Chomsky 1981), with the central principle being that of phonological licensing.

(1) *Principle of Phonological Licensing (Kaye 1990, Harris 1994)*

a. Within a domain all phonological units must be licensed except one, the head of that domain.

b. Licensing relations are local and directional.

In other approaches (e.g., Ito 1986), phonological licensing is conferred by virtue of units being incorporated into the prosodic hierarchy. In GP, however, licensing depends on a syntagmatic relation between adjacent positions. Heads do not need to be licensed within their particular domain. The dependent of a branching onset, nucleus or rhyme, however, requires licensing by an adjacent head position. The path of intraconstituent licensing is illustrated by the arrow between positions in Figure 1.1 below. Intraconstituent licensing is what allows constituents to branch, so it legitimizes branching onsets, nuclei and rhymes.

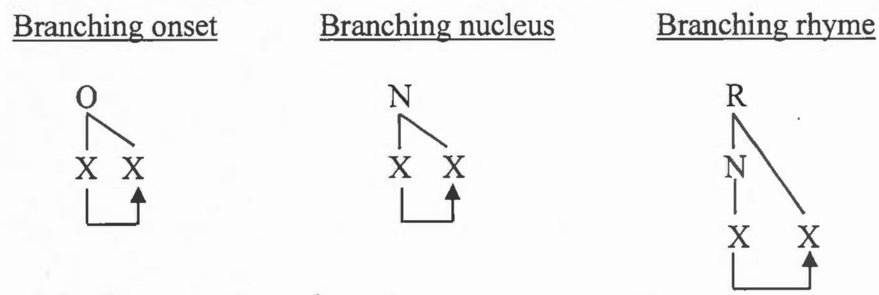


Figure 1.1 *Intraconstituent licensing*

Licensing is strictly local – that is, it operates between positions that are adjacent at some level of projection. Licensing is also strictly directional – in terms of intraconstituent licensing, this means that it operates exclusively from left to right. As a consequence of these two conditions, syllabic constituents are *maximally binary.* Three-member onsets, for example, are ruled out, since regardless of where the head is placed, licensing is violated either in terms of locality or directionality. As shown in Figure 1.2, if the head is at the left (or right) edge, locality is violated; if the head is in the middle, directionality is violated.

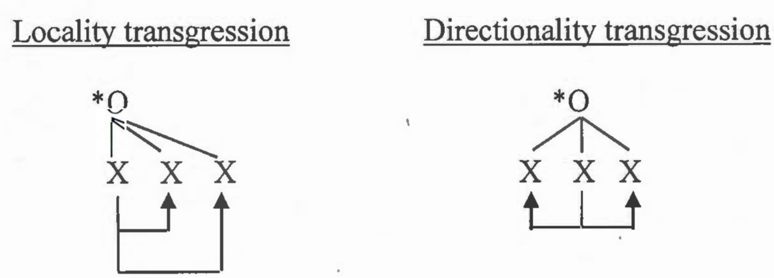


Figure 1.2 *Locality and directionality transgression*

Licensing operates not only between the members of a branching constituent but also between constituents. For example, a nucleus, as ultimate head of the syllable, licenses a preceding onset head. Between syllables, an onset licenses a preceding coda – see Figure 1.3. Unlike intraconstituent licensing, then, interconstituent licensing operates from right to left.

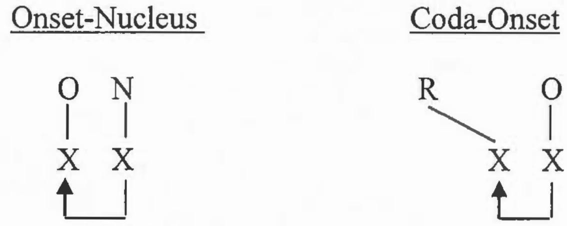


Figure 1.3 *Interconstituent licensing*

Codas are not constituents per se in GP. Instead, post-nuclear consonants are viewed as dependents of the rhyme or 'rhymal adjuncts', as shown. Still, the familiar term 'coda' continues to be used. Codas/rhymal adjuncts do not branch since, as pure dependents, they cannot license a dependent position of their own. GP thus limits syllabic constituents to the following: onset, nucleus and rhyme. The syllable node itself is not recognized in GP; rather, the rhyme projects up directly to the foot level.[4]

Certain forms of licensing involve the stricter relation of government. These include the intraconstituent relations between an onset head and its dependent, and between a nuclear head and its dependent, as well as the interconstituent relation between an

---

[4] I will adopt the GP practice of leaving out the syllable node, partly because it is of no consequence for the issues I address. I recognize, nonetheless, that there are arguments in favour of a syllable node (and I will continue to use the term 'syllable', along with 'coda', informally). For example, certain phenomena seem to be more easily accounted for by reference to the syllable, and specifically to stressed versus unstressed syllables. Notably, glottal stop insertion in standard German is sensitive to the stress status of the syllable: glottal stops are inserted before vowel-initial stressed but not unstressed syllables (Wiese 2000; Alber 2001). Also, in the process of *accent d'insistance*, which creates special emphasis in French, normal word-final stress shifts to the beginning of a word. However, in the case of *accent d'insistance dit affectif* (a particularly strong emphasis), stress shifts to the first syllable of a word only if it starts with a consonant; otherwise the second syllable is stressed. Hence: *formidable*, but *imbécile* (Gendron 1984). That is, the process is sensitive to whether the first syllable has a filled onset. In both German and French, then, the following condition applies: the onset of the head syllable of a foot must be filled. This is harder to express if the syllable constituent is absent and onsets are not projected to a higher node (but cf. Harris 1997).

Moreover, in some languages the tone-bearing unit is not the nucleus but apparently the entire syllable. Again, this phenomenon is easier to capture with reference to a syllable node. Likewise, certain forms of reduplication seem to target the syllable as a unit. Finally, language games such as *verlan* in French are often better captured by reference to the syllable. In sum, GP may have eliminated a valid level of the prosodic hierarchy. All the same, since the syllable node is not required for the issues addressed here, in the interest of simplicity, I will adhere to the practice of leaving it out.

onset and the coda it licenses. A government relation places restrictions on the melodic content of the licensor and licensee positions. Initially, these restrictions were expressed in terms of the 'charm' value of the elements comprising a segment – a governor was obliged to have *positive* and its governee *negative* charm. This notion has been superseded in GP work, however, by the notion of complexity (Kaye, Lowenstamm & Vergnaud 1990; Harris 1990, 1994). A governor cannot be less complex than its governee, with complexity being determined by the number of elements in a segment's composition.[5] See (2) for a formal expression of this notion.

(2) *Complexity Condition (Harris 1990)*

> Let α and β be melodic expressions occupying positions A and B respectively. Then, if A governs B, β is no more complex than α.

To calculate complexity, GP envisions the melodic make-up of segments in terms of privative elements rather than SPE-based distinctive features. Various versions of the GP element system have been proposed (including Harris & Lindsey 1995, 2000, Ingleby & Brockhaus 2002, and Backley 2011), so there is no absolute consensus.[6] Still, the elements that typically define vowels are: I U A @. The following elements are also typically implicated in consonantal melody: ? h R N H L. Despite each element being associated more commonly with either vocalic or consonantal melody, many of the elements can occupy both nuclear and non-nuclear positions. Unlike distinctive features, most elements can also be individually realized. That is, a single element can stand alone in an expression. For example, the element U in a nuclear position is realized as [u]; R on its own constitutes a tap [ɾ], and so on.[7]

---

[5] Essentially, complexity replaces the more familiar notion of a sonority differential between segments.

[6] Element Theory is introduced only briefly here. The reader is directed to Chapter II for a more thorough discussion.

[7] I have adopted the GP custom of always using square brackets around IPA symbols (or else nothing at all), rather than slashes. GP considers that there is no substantive difference between underlying and surface forms. That is, they are made up of the same elements and are hence both potentially

Like distinctive features, elements can be combined. A crucial difference between elements and distinctive features is that elements are privative rather than bivalent; traditionally, SPE-based features have been viewed as bivalent. The monovalency of elements entails that they are either present or absent in an expression, whereas a bivalent feature is always present with either a plus or a minus value. Crucially, since elements are privative, complexity can be determined in terms of the number of elements that make up an expression. This calculation of complexity is not possible in systems using bivalent features.

As an example of how complexity and government work, we can consider branching onsets in English. Remembering that the onset head to the left governs the dependent to the right, we can look at some possible combinations. For example, a stop such as [k] (composed of the four elements h-@-?-H) can govern the liquids [r] or [l], (composed respectively of the two elements I-R and R-?). A stop can also govern the glides [j, w] (with the single elements I and U). A fricative such as [f] (with the three elements h-U-H) can also govern liquids (two elements) and glides (one element). In each case, there is a complexity differential between the head and the governed dependent. Significantly, the fricative [h] (composed of the single element h) cannot combine with liquids, although it can with [j] and in some dialects with [w] (also composed of a single element).[8] To recap, only if a segment α contains an equal or greater number of elements than a segment β may α govern β; if β has more elements than α, α cannot govern β.

Like dependents of branching onsets, codas are also required to instantiate a complexity differential with the adjacent onset licensor. Codas, however, employ special measures to conform to the differential. For example, in so-called 'Prince'

phonetically realizable. Consequently, there is no attempt to systematically distinguish the two levels of representation.

[8] This situation is by no means idiosyncratic to English: in Onandaga (Michelson 1986), neither [h] nor [?] can form the head of a branching onset, and both are single element consonants.

languages (due to Prince 1984), the content of a coda is partially or wholly determined by the onset: only homorganic nasals or full geminates are permitted in coda position. In other cases, the complexity of coda consonants may be reduced through loss of laryngeal features (see Lombardi 1995b, as well as Ito 1986). Slavic and Germanic languages, for example, commonly exhibit obstruent devoicing in coda position, or else voicing may be determined via spreading from the adjacent onset.[9] A range of strategies, then, can be employed to minimize the number of elements supplied by the coda in a coda-onset sequence.

Cross-linguistically, sC-clusters have special status. One important consequence of the Complexity Condition is that s+stop clusters as in [st], [sk] and [sp] cannot be branching onsets, even when they are word-initial. Since [s] contains only three elements (h-R-H), it is insufficiently complex to govern a four-element stop. Conversely, due to directionality, [s] cannot be a dependent within the onset. In addition, due to maximal binarity of constituents, sCC-clusters such as [str] or [skl] cannot belong to a single onset. As a consequence of this and of other evidence, Kaye (1992) concludes that word-initial [s] in certain clusters is a coda licensed by an empty nucleus, as shown below in Figure 1.4.[10]

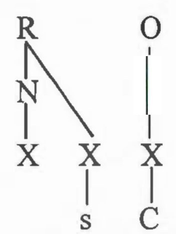


Figure 1.4 *Word-initial sC-cluster*

[9] See Brockhaus (1995), however, for an analysis of the context for devoicing as, in some cases, onset of an empty nucleus.

[10] Further support for this view comes from how sC-clusters are treated in second language acquisition. Learners without sC-clusters in their first language frequently insert a vowel into the initial empty nucleus (Carlisle 1991; Cardoso 2009). For example, Spanish speakers may pronounce English *student* as [e]*student*.

This is our first example of an empty category. Note that empty nuclei are not posited at random but as a consequence of a consonant sequence being otherwise unsyllabifiable. Note also that in GP empty nuclei always have the same representation: a nucleus dominating a timing slot with neither root node nor segmental content. The other representation I am arguing for, namely a nucleus with both timing slot and root node, is not used in GP. In the next section, the issue of empty nuclei in various positions in the word will be examined in greater detail.

1.1.1 Empty nuclei

Empty nuclei appear not only word-initially before sC-clusters; they also appear word-finally and word-internally. The presence of empty nuclei after word-final consonants follows from extensive evidence that such consonants cannot be codas (Kaye 1990; Harris 1997; Harris & Gussman 1998, 2002); instead, according to GP, they must be onsets. In addition, if they are onsets, they are necessarily followed by a nucleus for licensing purposes.

Briefly, one problem with the coda view of final consonants is that, in quantity-sensitive languages, where codas normally contribute weight for the purposes of stress assignment, final consonants sometimes fail to do so (Liberman & Prince 1977; Hayes 1981, 1982; Halle & Vergnaud 1987). A further enigma is that the set of word-internal coda consonants and the set of final singleton consonants frequently fail to match. In Diola Fogny and Pompanean (Ito 1986) and in English (Harris 1994), the set of final consonants is larger than the set of word-internal codas. Conversely, in languages such as Italian (Basbøll 1974), the set of final consonants is more restricted than the set of word-internal codas. In extreme examples, Luo, Yapese and Yucatec Maya allow any consonant in word-final position while word-internal codas are proscribed; Axininca Campa and Telugu have word-internal codas but no final consonants (Harris & Gussmann 2002). These distributional mismatches are not what one would expect if word-internal codas and final consonants have identical prosodic structure.

Issues such as these have led phonologists to propose alternatives to the coda view. One position is that a word-final consonant may be analyzed as an extraprosodic appendix (Halle & Vergnaud 1980; Charette 1985; Vaux & Wolfe 2009) as in (7).

Figure 1.5 *Extraprosodic appendix*

Since it is not associated to a syllable constituent, the appendix in Figure 1.5 is licensed instead at a higher level of structure such as the prosodic word. Its status as appendix, unattached to any syllabic node, is thought to explain the failure of a final consonant to contribute to syllable weight. The discrepancies in distribution between final consonants and internal codas are also attributed to this distinction in representation.

Where the appendix analysis runs into trouble is with final consonant sequences, since it fails to predict the attested patterns. If a word-final appendix truly bypasses the syllable level of prosodic structure, it should circumvent the complexity (or sonority) restrictions that normally apply within and between syllables. But this is not what we find. Word-final consonant sequences are generally required to show rising complexity, precisely the pattern that applies word-internally to coda-onset sequences. In English, rising complexity pairs of consonants such as [rt] and [lm] appear word-finally, but falling complexity *[tr] and *[ml] do not.

In French, at first blush, the rightmost consonant in a word-final pair looks immune to complexity restrictions, since both rising and falling complexity clusters are attested: for instance, both [rt] as in *porte* and [tr] as in *vitre* are found. The freedom is only partial, however: in a word-final pair with falling complexity, only liquids are permitted in second position; sequences such as *[tn] or *[tf] are thus ruled out. In effect, final sequences with falling complexity are limited to combinations that constitute a well-formed branching onset elsewhere. In brief, if final consonants have the status of extrasyllabic appendix, the distributional restrictions that apply in English and French are hard to explain.

Another alternative to the coda view is that final consonants are onsets. This position follows from the GP principle of coda licensing (Kaye 1990). [11]

(3) *Principle of 'Coda' Licensing*

Post-nuclear rhymal positions must be licensed by a following onset.

This principle strictly bars word-final consonants from being codas, since final consonants are not followed by an onset. If final consonants are onsets, however, an adjacent nucleus is required for licensing purposes: according to interconstituent licensing, onsets need to be licensed by a nucleus. In the case of final consonants, of course, this nucleus is phonetically unexpressed.

Final empty nuclei are posited not only for the purpose of licensing the preceding onset. More concrete evidence for the presence of an empty nucleus comes from vowel-zero alternations: in some languages, at times the nucleus remains empty, but

[11] See, however, Piggott (1999) and Rice (2003) for evidence that challenges the universality of coda licensing. I return to this issue in Chapter V.

at others it is realized. Realization of an empty category can generally take one of two forms: either a default segment is inserted or else a segment is determined by spreading from a local source. With empty nuclei, the default segment is often a schwa; otherwise, spreading from an adjacent nucleus results in an identical 'copy' or else a harmonic vowel.

For example, empty nuclei after consonant-final words in French are not normally realized. However, under certain conditions, a final schwa may be produced. Notably, when the initial member of a compound ends in two or more consonants, a final schwa is realized, as long as the second word is (phonetically) monosyllabic.

(4) *Final empty nuclei in French (Charette 1991)*

Without schwa		With schwa	
port(e)	'carry'	*porte-clefs*	'key ring'
gard(e)	'keep'	*garde-fou*	'railing'
couvr(e)	'cover'	*couvre-lit*	'coverlet'
ouvr(e)	'open'	*ouvre-boîte*	'can opener'

Similarly, certain forms in Selayarese end in an empty nucleus that exhibits vowel-zero alternation (Mithun & Basri 1986). Rather than being a default segment as in French, however, the vowel is determined by spreading from the preceding nucleus. Some examples follow (accents indicate stress).

(5) *Final empty nuclei in Selayarese*

sá:hala	'profit'	sússulu	'burn'
lámbere	'long'	kí:kiri	'metal file'
bé:rasa	'rice'	tú:lisi	'write'

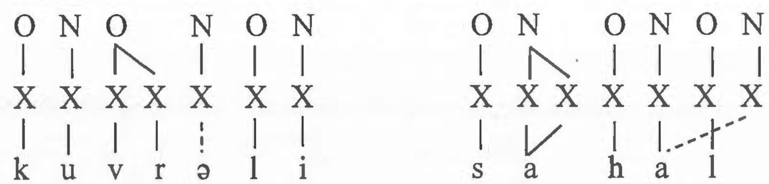
Briefly, evidence that the final nuclei in these forms are lexically empty comes partly from their stress patterns. Normally, stress in Selayarese is penultimate. However, in the forms above, stress is antepenultimate. Compare [sahá:la] 'sea cucumber', which has a final fixed vowel and normal stress, with [sá:hala] 'profit', which has a final copy vowel and penultimate stress. Apparently, for purposes of stress assignment, the final vowels in these forms do not count, whereas usually final vowels do count: [sá:hal(a)] 'profit' versus [sahá:la] 'sea cucumber'. Additionally, the final vowels in the forms above alternate with zero before a vowel-initial suffix, whereas final vowels in other forms do not. Compare [lámbere] 'long' and [lámber-aŋ] 'longer' with [lóhe] 'many' and [lóhe-aŋ] 'more'. In sum, the forms in (10) end in an empty nucleus that is at times filled by spreading from the preceding vowel.

The case of final vowels in Selayarese provides a simple illustration of how representational distinctions can clarify differences in behaviour of apparently identical forms. Although the vowels at the end of [lámbere] and [lóhe] are to all appearances identical, the former [e] is realized in an empty nucleus via spreading whereas the latter is an underlying [e]. As a consequence, only the [e] at the end of [lámbere] alternates with zero.

The two ways of realizing an empty nucleus are illustrated provisionally below. For a default vowel, melody is inserted into the empty nucleus (6a); for a copy vowel, the melody in one nucleus spreads to another, creating a doubly linked structure (6b).

(6) *Realization of final empty nuclei*

a. Default vowel 'couvre-lit'   b. Copy vowel [sá:hala]



Empty nuclei are not limited to initial and final sites. Word-internally, empty nuclei are posited when a sequence of consonants cannot be parsed as either a branching onset or a coda-onset sequence. For example, the English words *kidney* and *aṭlas* are analyzed as [kɪ.dØ.ni] and [æ.tØ.lə.sØ] because the surface-level sequences [dn] and [tl] do not fit the profile of either well-formed complex onset or coda-onset. In reverse order, [nd] and [lt] would of course be canonical coda-onsets. Further evidence for word-internal empty nuclei comes again from vowel-zero alternations – see the examples from Quebec French and Khalkha Mongolian in (7) (based on Charette 1991).

(7) *Internal empty nuclei*

Quebec French			Khalkha Mongolian		
s(e)mer	[sØ.me]	'to sow'	*am_raa*	[a.mØ.ra:]	'one's lover'
parsemer	[par.sə.me]	'to sprinkle'	*amar*	[a.ma.rØ]	'lover'

GP makes no representational distinction between an empty nucleus that receives no realization, as in *port(e)* or *s(e)mer,* and an empty nucleus that is realized by a default or copy vowel, as in *porte-clefs* and *parsemer* or *sá:hala* and *amar.* Instead, one particularity of GP is that empty nuclei need to be licensed in order to remain empty (i.e., in order not to be phonetically expressed). In the absence of licensing, empty nuclei obligatorily receive phonetic expression. Vowel-zero alternations then are due to empty nuclei being licensed or not, depending on the context.

A different way of handling vowel-zero alternations, however, might be to vary the representation of the empty nucleus. When it contains a timing slot and root node, but no lexically specified segmental content, a process is automatically triggered to fill the empty root node (8a). When the empty nucleus contains only a timing slot, however, no such process is triggered (8b). This is not the standard GP approach to vowel-zero alternations. Instead, GP attributes such variations to differences in licensing of the empty nucleus. In cases such as (8a,b), of course, both the licensing and representational approaches can potentially account for the alternations, since the contexts are not identical.

(8) a. *porte-clefs* b. *port(e)*

Indeed, Charette (1990) proposes that the empty nucleus in *porte* is only licensed in domain-final contexts; hence, the empty nucleus in the compound *porte-clefs* is realized because it is domain-internal. In Charette's view, final unrealized nuclei confer a 'license to govern' on a preceding onset, permitting it to govern a dependent coda; however, internal empty nuclei must be realized in order to act as government licensor. In a system using a representational distinction, it could be argued the empty nucleus without a root node is a government licensor only in domain-final sites.[12]

One situation that would be problematic for the GP licensing approach would be cases of empty nuclei in the same context, some of which remain empty and others of which are realized. From a representational perspective, on the other hand, differences in the behaviour of empty nuclei in identical contexts can be attributed to differences in degree of degeneracy of the representation. In Chapter IV, we will examine instances of precisely this situation involving empty nuclei. In a nutshell, this is the type of problem to which the representational approach provides a solution.

Before we consider such problems, we need a clearer picture of exactly how licensing of empty nuclei operates in the GP framework. Depending on where the empty nucleus appears in a word, different types of licensing apply. Word-final empty nuclei are licensed by parameter. In languages where the parameter is set at 'off' (e.g.,

[12] Note that this is not a rejection of the notion of 'license to govern' per se, but merely a reformulation of it that links the ability to confer a license to govern to the degree of degeneracy of the domain-internal empty nucleus.

Hawaiian, Italian or Japanese[13]), final nuclei are always filled, whether with a full or default vowel or via spreading from an adjacent position. In this case, words never end in a consonant.[14] In languages with the parameter set at 'on' (e.g., English, French or Khalkha Mongolian), final nuclei may be empty and remain phonetically unexpressed.

Word-initially, licensed empty nuclei are specific to sequences starting with [s]. Empty nuclei before sC-clusters are said to be licensed by 'magic' (Kaye 1992). The term captures the rather mysterious requirement that [s] alone appear in this context.[15] Essentially, Kaye proposed the term as a stop-gap measure to indicate that we do not fully comprehend the mechanism by which such empty nuclei are licensed. Provisionally, magic licensing also constitutes a parameter that may be set at 'on' or 'off' in a language.

Word-internal empty nuclei differ in being licensed either by 'proper government' or, more rarely, by 'interonset government'. Essentially, proper government entails that an empty nucleus is sanctioned to remain empty by an adjacent nucleus that is itself phonetically expressed. (9) contains a formal expression of proper government.

(9) *Proper Government (Kaye 1990)*

A nuclear position α properly governs a nuclear position β iff:

a) α is adjacent to β on its projection
b) α is not itself licensed [*i.e., to remain empty*]
c) no governing domain separates α from β

---

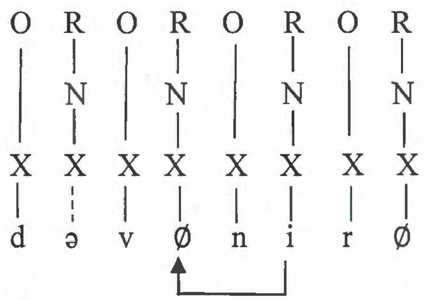
[13] The presence of final nasals in Japanese could be taken as evidence for the parameter being set at 'on', but see Yoshida (1991 – cited in Honeybone 1999) for arguments against such an analysis.

[14] In the initial stage of second language acquisition, this setting is transferred from the first to the target language. The transferred setting results in the production of a vowel at the end of consonant-final words, a feature of pronunciation characteristic of, for example, Italian learners of English.

[15] Although this is not addressed in standard works of GP, empty nuclei in fact also appear before initial full and partial geminates, as in Selayarese [ppa:lui] and [nta:le] (Piggott 2003b). This situation will be examined in forthcoming Chapter IV.

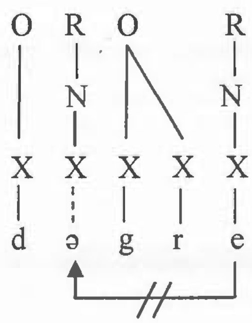
Depending on the language, proper government may operate from right to left (the more common choice) or from left to right. To illustrate, French *devenir* ‘to become’ contains two word-internal empty nuclei, one in the first and another in the second syllable. As indicated in (10) by the arrow showing the flow of licensing, the third nucleus (being expressed) properly governs the preceding empty nucleus, permitting it to remain empty. By virtue of being phonetically unexpressed, however, this same nucleus cannot license the preceding empty nucleus to remain empty. Hence the first empty nucleus is realized by a default schwa. Clauses a) and b) of proper government are thus demonstrated, and the resultant surface form is [dəvØnirØ].

(10) French *devenir:* dØvØnirØ → dəvØnirØ



Clause c) states that proper government will fail to apply if a governing domain intervenes between the empty nucleus and its licensor. The kind of governing domain that might intervene is either a branching onset, as with the example from French in (11), or else a coda-onset sequence.

(11) French *degré:* dØgre → dəgre

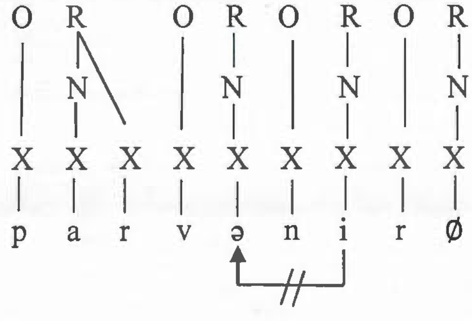


The empty nucleus in the first syllable of *degré* 'degree' fails to be properly governed by the following nucleus due to an intervening branching onset. Consequently the empty nucleus must be phonetically expressed: [dəgre], not *[dØgre].

Examples of coda-onset sequences preceded by an empty nucleus are more rare, presumably because an empty nucleus, whether realized or not, makes for a poor licensor of the coda. To my knowledge, they do not occur in French (with the exception of initial sC-clusters preceded by an empty nucleus); in English, they appear in forms such as *complete* [kəm.plitØ], *perform* [pərfɔrmØ], and *illustrate* [ɪləstret]. In each case, the intervening governing domain blocks proper government from applying to the empty nucleus.

Earlier we touched on a further restriction to the application of proper government proposed in Charette (1990). In languages such as French, no governing domain may *precede* the properly governed position. That is, an empty nucleus following a branching onset or a coda-onset sequence cannot be properly governed and must be realized. For example, the empty nucleus in the second syllable in *parvenir* 'to attain', unlike *devenir*, is produced with a schwa: [parvənirØ]. Due to the [r.v] coda-onset sequence, the empty nucleus cannot remain empty, despite the presence of the potential proper governor [i] in the third syllable. In Charette's terms, the onset requires a 'license to govern' the coda; this license can only be supplied by an expressed nucleus. In instances where a license to govern is required, then, proper government fails to apply, as shown in (12).

(12) French *parvenir:* parvØnirØ → parvənirØ



At other times, however, word-internal empty nuclei have been found to remain empty despite the conditions for proper government not being met. These cases have been explained by the notion of interonset government. With interonset government, an empty nucleus is licensed by a governing relation between flanking onsets. This mechanism is invoked notably when sequences of adjacent empty nuclei remain unexpressed. More than one empty nucleus in a row should be illicit, since an empty nucleus needs to be licensed by a following realized nucleus. According to proper government, then, at most every second empty nucleus should remain empty.

Interonset licensing of word-internal empty nuclei was proposed by Gussmann and Kaye (1993) to account for sequences of three or more consonants at the beginning of words in Polish (see Cyran 1996 for the application to Irish).

(13) *Initial clusters in Polish*

pchła	[pxwa]	'flea'
mgła	[mgwa]	'mist'
mgnienie	[mgnɛnɛ]	'batting of eyelid'
cknić	[t͜sknit͜ɕ]	'hanker'
mdleć	[mdlɛt͜ɕ]	'faint'

Polish permits branching onsets comprised of obstruents and liquids/glides such as [bl, dr, kl, tw, gw]. However, the initial clusters in (13) do not have a branching-onset profile. First, from a traditional sonority-based perspective, the sequences are suspect since both sonority rise [pxw] and fall-rise [mgw] are found. Next, from a purely GP perspective, the condition of maximal binarity imposed on constituents categorically excludes such sequences from constituting branching onsets.

Regarding [pxw] and [mgw], it might be possible (albeit mistaken) to analyze the sequences as a simplex onset followed by an empty nucleus and then a branching onset: that is, as [pØ.xwa], [mØ.gwa]. In this case, we would only have to explain why the nucleus [a] is able to properly govern the empty nucleus across the branching-onset governing domain. Of course, such an analysis could not apply to the other forms, since [gn], [kn] and [dl] are illicit branching onsets. In addition, even the potential branching onsets [xw] and [gw] are dubious, given the existence of vowel-zero alternations – see the genitive plural forms with a vowel in (14).

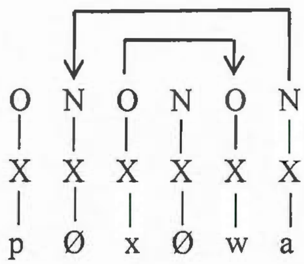
(14) *Vowel-zero alternations in initial clusters in Polish*

Nom. sg.		Gen. pl.	
pchła	[pxwa]	*pcheł*	[pxɛw]
mgła	[mgwa]	*mgieł*	[mgɛw]

Given these alternations, it seems the best analysis for these and all the other words in (13) above is to view them as sequences of three separate onsets with two intervening empty nuclei: that is, as [pØxØwa], [mØgØwa], [mØgØnɛnɛ], and so on.

The problem with this scenario is that the two empty nuclei cannot both be properly governed. This led Gussmann and Kaye to propose that that second and third consonants in each of the above sequences are in an interonset relationship. By virtue of this relationship, the intervening empty nucleus is licensed. Interestingly, a condition seems to apply to such consonants: only a sequence of obstruent followed by sonorant can form an interonset pair. In effect, all the Polish pairs resemble coda-onset sequences in reverse. Apparently, proper government can operate over the interonset governing domain. As a result, the first empty nucleus in a form like [pØxØwa] is licensed to remain empty by proper government from the third nucleus; meanwhile the second empty nucleus is licensed by interonset government (15).

(15) *Interonset and proper government of empty nuclei*



Like 'magic' licensing, interonset licensing remains a bit of a mystery, and it carries with it an unfortunate impression of the stipulative. In phonological systems, interonset licensing is a relatively rare phenomenon, which invites the question of why it is not universally manifested. A further question that remains unanswered concerns why the interonset relation results in licensing of the empty nucleus, especially since normally it is a nucleus that licenses an onset.[16]

In sum, four ways of licensing empty nuclei have been proposed in GP: i) domain-finally, by parameter; ii) domain-initially, before sC-clusters, by 'magic' parameter; and domain-internally, by either iii) proper government or iv) interonset government. Empty nuclei that are not licensed to remain empty must be realized. As we have seen, the quality of the vowel that appears in an unlicensed empty nucleus varies. In some languages, a default vowel such as a schwa is produced; in other languages, the vowel quality is determined by the context, through spreading from an adjacent vowel (or, more rarely, from a consonant). Given that onsets and codas need to be licensed in order to support melodic content, it may of course seem counter-intuitive that empty nuclei require licensing in order to remain devoid of content. This paradox counts as one argument against the GP approach involving licensing of empty nuclei. Another feature that disfavours the approach is the sheer quantity and complexity of the various means for licensing empty nuclei. As stressed earlier, however, the crucial problem that the representational approach solves concerns empty nuclei behaving

[16] I am grateful to Mohamed Guerssel (personal communication) for these critical observations.

differently in different items within one and the same context, some nuclei being licensed to remain empty, others requiring realization. The forthcoming Chapter IV on empty nuclei will present some crucial evidence of precisely this kind of behaviour. In the meantime, we address the question of how empty onsets and empty codas are represented in GP.

1.1.2 Empty onsets

Overall, less emphasis has been placed in GP on empty onsets, perhaps because their presence is not as contentious as empty nuclei in mainstream phonological theory. Empty onsets are similar to empty nuclei in that they may at times remain empty and at others receive phonetic expression. One difference with empty nuclei is that there is no mechanism for licensing empty onsets to remain empty. Indeed, there is a general preference for empty onsets to be phonetically realized in some way. Like empty nuclei, not all languages permit empty onsets, but languages that do permit them employ various processes to fill the empty onset. These include glide formation in hiatus, sandhi linking phenomena, and default consonant insertion (typically of a glottal stop). In GP, these onset-filling processes constitute evidence for the presence of an empty onset in a representation. Hence, onsets are not inserted to fulfill phonological requirements; they are already present and then filled by various means.

Theory-internal justification for empty onsets comes from the notion of head-dependent asymmetry: heads are obligatory, whereas dependents are optional. Since onsets are heads, these are obligatory constituents. Consequently, in the GP view, wherever a nucleus appears in a representation, there must be a corresponding onset, even if this onset remains empty (but, for an opposing view under which onsets may at times be absent, see Guerssel to appear). A pair of words such as *cold* and *old,* then, are distinguished not by the presence versus absence of an initial onset, but by the presence of a filled versus an empty onset. In English, this empty onset may at times remain empty as in [_oul.dØ]. More commonly, however, it is filled, whether by a glide as in *the old* [ði.joul.dØ], by a linking consonant as in *an old* [ə.noul.dØ], or, when neither of these options is available, by a default glottal stop as in *old* [ʔoul.dØ].

When a glide is formed in hiatus position in *the old*, the vowel in *the* (or more precisely, the element I) spreads into the empty onset, creating a doubly linked segment (16a). When the [n] in *an* links up with the beginning of the word *old*, the segment delinks from its original position and re-associates to the empty onset (16b). Finally, when a glottal stop is realized at the beginning of *old*, default content (the element ʔ) is inserted into the empty onset (16c).

(16) *Realization of empty onsets*

```
a.  O N O N          b.  O N O N          c.  O N
    | | | |              | | | |              | |
    X X X X              X X X X              X X
    | |/  |              +/    |              : |
   [ð i   o u l d]     [æ n    o u l d]      [ʔ o u l d]
```

Interestingly, some languages seem to have more than one type of empty onset. That is, one set of empty onsets behaves differently from another. This is precisely the kind of problematic difference in behaviour of apparently identical forms that is best accounted for by a representational distinction. One example comes from French, which has two types of non-consonant-initial words: a conventional set and a special *h-aspiré* set. In terms of melodic content, both types of words are lexically vowel-initial, so they both start with an empty onset. Words from the *h-aspiré* set, however, generally behave as though they were consonant-initial (i.e., with a filled onset). This behaviour is attested in processes such as liaison and elision of the vowel in the definite article *le/la*. Some minimal pairs involving *h-aspiré* versus truly vowel-initial words are provided below in (17). Note that the *h-aspiré* set are all orthographically h-initial, corresponding to a historical [h] (but not all h-initial words are *h-aspiré*).

(17) *Two types of empty onsets in French*

Vowel-initial words		*H-aspiré* words	
aine	'groin'	*haine*	'hate'
ache	'wild celery'	*hache*	'axe'
une	'one'	*hune*	'top'
ère	'era'	*hère*	'sorry fellow'

The difference between the two types of words can be shown by how the definite article behaves preceding *aine* and *haine*: *la* + *aine/haine.* With vowel-initial words such as *aine,* the definite article loses its vowel (*la* → *l*). In this case, [l] is able to fill the empty onset position. The pronunciation of *la* + *aine* is thus [lɛnØ]. With *h-aspiré* words such as *haine,* on the other hand, the definite article maintains its vowel. The pronunciation of *la* + *haine* is thus [la_ɛnØ], and the empty onset remains empty (if it is not filled by a default glottal stop). The challenge is how to capture this distinction representationally, since both sets of words appear to start with an empty onset.

Conventionally in GP, as shown in representations so far, empty nuclei are represented as having a timing slot and no melodic material. For empty onsets, on the other hand, two representations have been proposed: an onset with neither melody nor timing slot, and an onset without melody but with a timing slot (Charette 1991). The former representation usually applies to empty onsets before vowel-initial words; the latter is reserved for special cases such as *h-aspiré* words in French – see Figure 1.6.

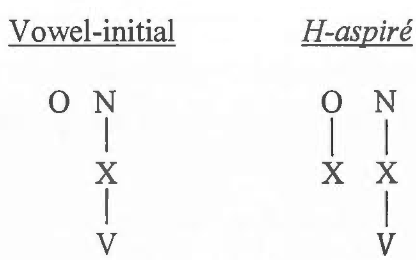


Figure 1.6 *Two representations for empty onsets*

This is exactly the type of abstract representational distinction that I am advocating to deal with differences in the behaviour of apparently identical forms. All the same, as will be argued in Chapter III, the distinction that has been adopted in GP for *h-aspiré* versus vowel-initial forms needs to be revised in favour of a distinction based on presence versus absence of a root node in the empty onset.

In sum, word-initially, empty onsets appear in forms that start phonetically with a vowel (e.g., ØVCV); word-internally, they appear in hiatus position between heterosyllabic vowels (e.g., CVØV). From the GP point of view, syllables without an initial consonant do not simply lack an onset; they have an onset that is lexically empty. Wherever there is a nucleus, an onset must also be present, since onsets are heads and heads are obligatory constituents. The presence of this empty onset is confirmed by a range of processes that serve to provide the position with content.

1.1.3 Empty codas

While empty onsets have received less attention than empty nuclei in GP, empty codas are truly marginal. Although they are mentioned in Kaye, Lowenstamm and Vergnaud (1990), to my knowledge empty codas have never been explored in any detail. An empty coda constitutes the first half of a geminate consonant. By extension, a partial geminate (i.e., a homorganic nasal + obstruent) involves a partially empty coda, one that is specified minimally for the element N.[17] The coda specified only for nasalization can be considered partly empty since N is one of the few elements that cannot receive stand-alone realization. Some examples of full and partial geminates from Japanese are given below in (18).

[17] Harris (1994) represents partial geminates as lexically specified for both N and ?. In Chapter V, I argue that only N needs to be present in the lexical representation.

(18) *Japanese geminates*

Full geminates		Partial geminates	
kappa	'river'	tombo	'dragonfly'
kitta	'cut, sliced'	kande	'chewing'
sakka	'author'	kaŋgae	'thought'
kassen	'battle'	unzari	'disgusted'

Kaye, Lowenstamm and Vergnaud (1990) are quite explicit in representing geminates as involving an empty coda. A geminate is formed by spreading the content of an onset into a preceding empty coda, as shown in Figure 1.7.

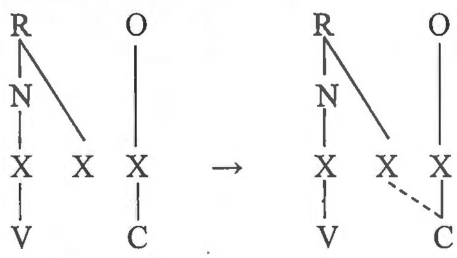


Figure 1.7 *Geminate formation*

In the formation of a geminate, the empty coda receives all of its content from the adjacent onset. Similarly, with partial geminates the coda is lexically specified for nasality and receives its place specification from the neighbouring onset (Figure 1.8).

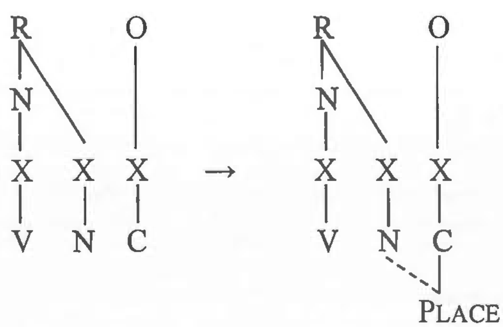


Figure 1.8 *Partial geminate formation*

Normally there are two ways for empty categories to signal their presence in a representation: i) because they are required for licensing purposes or ii) because they at times appear as filled. With empty codas, however, only the latter applies. Codas are not actual constituents in GP but 'rhymal adjuncts'. Unlike nuclei and onsets, then, codas are pure dependents and thus have no licensing powers of their own. Consequently, empty codas are only required in a representation if they are realized.

When empty onsets are realized, they may receive their content by default or by spreading. In Japanese, empty codas can only be filled from a local source: the neighbouring onset. In Chapter V, however, it will be argued that empty codas in Selayarese may also receive default content in the form of a glottal stop, as shown below in Figure 1.9.

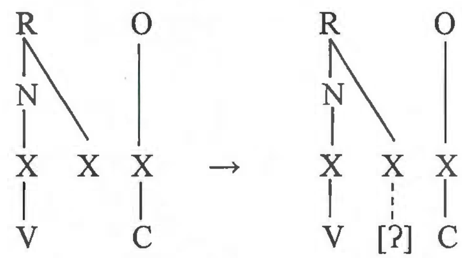


Figure 1.9 *Default consonant in an empty coda*

Selayarese in fact has full gemination alongside default glottal stop in empty codas, and the two strategies for filling the empty coda are instantiated in the same contexts. This is problematic if the empty coda simply has the same representation in both cases. What is it that determines whether the coda is filled by spreading from an adjacent onset or by default insertion of content? Again, the solution that will be proposed is that we are dealing with empty categories of differing degrees of degeneracy: when the empty coda contains a root node, it triggers default insertion of glottal stop; when it lacks a root node, it receives all of its content from the root node under the adjacent onset, generating a geminate consonant.

As with empty nuclei and onsets, not all languages permit empty codas. Hence, despite Prince's (1984) assertion that full and partial geminates are the unmarked option for codas, the situation is not so simple. While it appears to be true that all languages with codas permit homorganic nasals in coda position, not all languages with codas permit full geminates. Essentially, the presence of gemination in a language appears to be entirely independent of the ability to support filled codas. Japanese allows full and partial gemination, but no other coda consonants; English allows a whole range of coda obstruents and sonorants (including homorganic nasals), but not geminates. This kind of distinction between languages is usually handled by parameter in GP. I propose the following.

(19) *Empty coda parameter*

Codas may be empty.

The presence versus absence of geminates in a language is thus not accidental: it depends on a parametric option. In Japanese, the empty coda parameter is set at 'on', such that gemination is permitted; in English, the parameter is set at 'off', such that gemination is excluded. Because partial gemination involves only partly empty codas, apparently it is permitted even if the empty coda parameter is set at 'off'.

In sum, codas may be underlyingly empty, but they never remain empty – this option is limited to empty onsets and nuclei, whose presence is required independently for licensing purposes. In surface forms, empty codas are realized as the first half of a full or partial geminate or, more rarely, as a default consonant such as a glottal stop. Just as not all languages permit codas, only some languages permit empty codas, as a function of parametric licensing. Finally, as with empty nuclei and empty onsets, it will in fact prove necessary to posit two degrees of degeneracy for empty codas: these may appear either with or without a root node.

1.2 Proposed representations

In the preceding review of the GP framework, empty nuclei and empty codas are always represented the same way: as timing slots associated to a syllabic constituent and lexically devoid of segmental content. Empty onsets, on the other hand, are given two representations: as an onset constituent devoid of segmental content, but with or without a timing slot – see Figure 1.10.

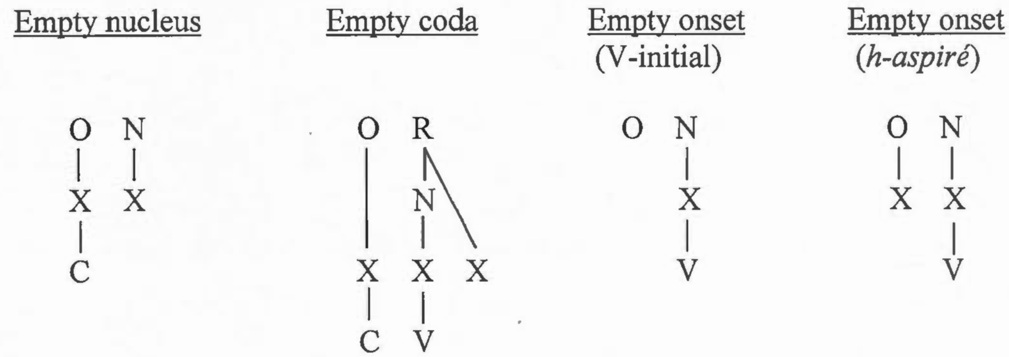


Figure 1.10 *GP representations of empty categories*

There are problems, however, with this set of representations. First, just as two different types of empty onsets are required to account for the full range of behaviour of non-consonant-initial forms, I will show that there is a need for more than one representation of empty nuclei and empty codas. The motivation comes partly from examples of empty nuclei and empty codas that behave differently in the same context. Second, the representational distinctions that have been employed to date for empty onsets are not the optimal solution. Instead, I will propose that there are two possible representations for all empty categories: i) containing a timing slot and root node; and ii) containing only a timing slot. These are shown below in Figure 1.11 with respect to empty nuclei, empty onsets and empty codas.

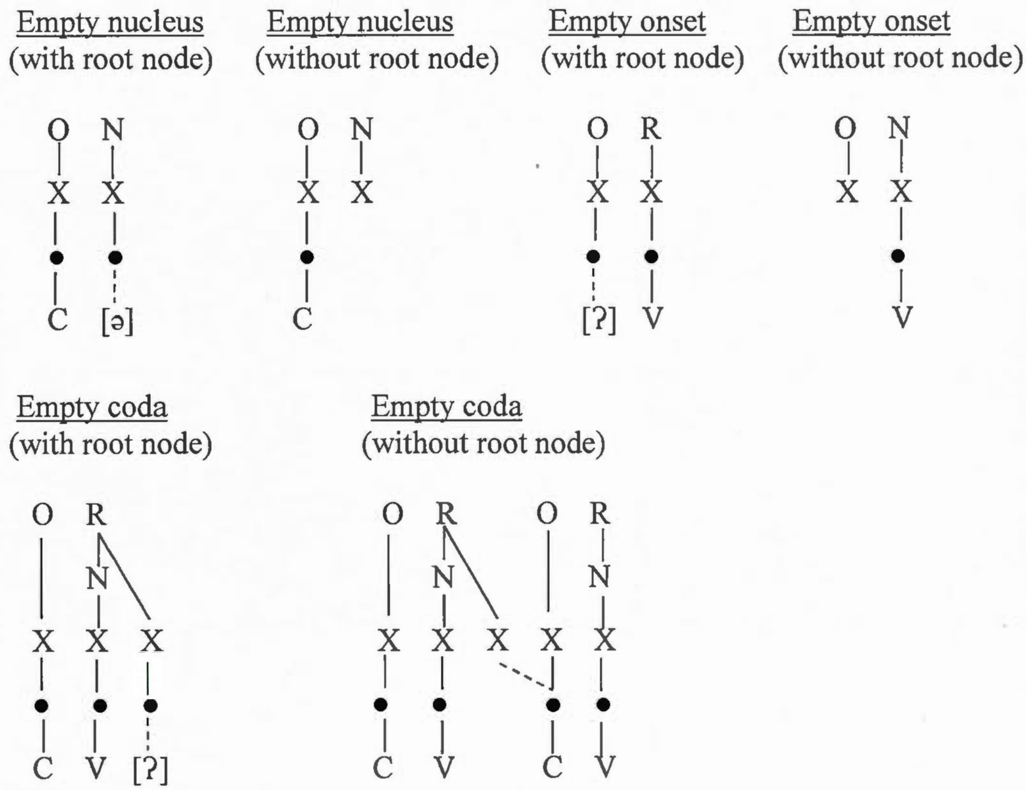


Figure 1.11 *Proposed representations of empty categories*

If a representational distinction between two types of empty category is indeed required, my proposal nonetheless invites the question: Why must this distinction be based on the presence versus absence of a root node? Is this not an arbitrary distinction that could just as well be captured via some other means (including the presence versus absence of a timing slot or indeed via lexical markings)?

Crucially, empty categories sometimes trigger production of a default segment (e.g., a schwa in an empty nucleus and a glottal stop in an empty onset or coda as shown above in Figure 1.11). My contention is that this phenomenon is incumbent on the presence of a root node: the sole function of this prosodic unit is to harbour melody. In the absence of lexically specified melody, then, it follows that the root node is capable of generating its own content. Conversely, it follows that empty nuclei and onsets without a root node are incapable of triggering default melody; they lack the requisite prosodic unit and hence tend to remain empty as also shown in Figure 1.11.

Aside from the argument from default content, convincing evidence for a distinction based on presence versus absence of a root node comes from processes of local spreading that target empty categories (e.g., vowel harmony, glide formation or gemination). Two types of spreading processes are found: i) spreading of individual elements from complex segments; and ii) spreading of an entire set of elements from a complex segment.[18] These are shown in Figure 1.12.

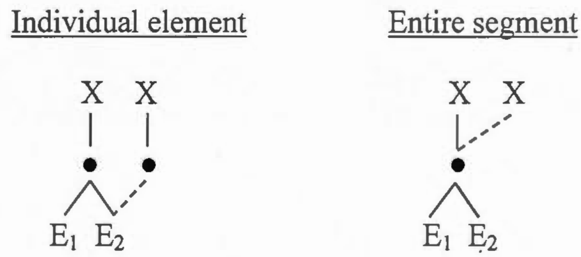


Figure 1.12 *Two types of spreading process*

The first type of process involves elements that are doubly linked to two root nodes: the root node in the source segment that harbours the element and the root node in the target. Necessarily, with this kind of spreading process, the target empty category must contain its own root node; otherwise the individual element that spreads would have nowhere to dock. The second type of process, conversely, involves a root node doubly linked to two timing slots. The presence of a root node under the target timing slot would thus be at best superfluous to spreading. Indeed, restriction of spreading to precisely those cases where all the elements under an adjacent root node are shared strongly implies the absence of a root node in the target. As we will see in greater detail in subsequent chapters, then, the type of local spreading that occurs in an empty category depends on its representation. The presence versus absence of a root node is

[18] When the source is a simplex segment, either type of spreading would generate the desired output.

thus not an arbitrary distinction; it crucially establishes the ability of the empty category to license default melody and determines the type of spreading process that is potentially available.

1.3 Distribution of empty categories

So far, I have presented empty categories (whether with or without root nodes) as simply representational options that can be employed lexically or that can emerge in the course of a derivation. In representations, in the same way that segments such as [p] versus [b] or [i] versus [u] can be used to distinguish lexical items, empty categories can also perform a distinctive function. Thus, in French, there is a contrast between words such as *laine* 'wool' that start with a filled onset and words such as *aine* 'groin' that start with an empty onset (in this case, without a root node). A contrast may also arise in French between empty onsets with and without a root node: for example, between an *h-aspiré* word such as *haine* 'hate' and a vowel-initial word such as *aine* 'groin'. That is, when languages use more than one type of empty nucleus, onset or coda, this distinction can be harnessed to contrast lexical items, in much the way that segmental contrasts are used.

Similarly, empty categories may be derived. Just as, say, [b] may become [p] in contexts that trigger devoicing, filled categories may become empty categories through loss of elements (leaving a bare root node) or through severing of the root node (leaving only a timing slot). When vowels in English reduce to schwa, for example, the process involves suppression of all element content, leaving only a nucleus with an unspecified root node that is filled by default content. When vowels are deleted, the process involves delinking of the root node, in the absence of which a default vowel cannot appear. Empty categories themselves may be subject to processes of fortition (insertion of a root node) and lenition (delinking of a root node). In this case, they are chameleons, at times exhibiting the behaviour of an empty category with a root node, at others without a root node. In a nutshell, root nodes can show the same changeability as timing slots, which may be suppressed (e.g., vowel-shortening) or added (vowel-lengthening).

The proposal to distinguish between empty categories with and without a root node provides invaluable insight into the workings of phonological systems. Nonetheless, I wonder whether it is not possible to be more ambitious and to circumscribe to some degree the distribution of empty categories. For example, are there contexts where empty categories are excluded or, conversely, where they are required? Are there contexts where an empty category must include a root node? That is, are there constraints on the distribution of empty categories and on their degree of degeneracy within a representation? On this point, some general observations can be made.

First, there is a tendency for empty categories, particularly onsets and nuclei, to be relegated to prosodically recessive positions. This holds doubly so for empty categories without a root node. For example, in some languages, empty nuclei (even with a root node) tend not to appear in stressed syllables – this is the case in English. Empty onsets that remain empty (i.e., that lack a root node) are also disfavoured in stressed syllables. In a sense, it follows from the Complexity Condition that empty categories should be banished from dominant positions, since empty categories have a complexity count of zero (or even less than zero, I suppose, if the root node is missing).[19] Next, relatedly, empty onsets and nuclei do not branch. For example, English has forms like *old* [ʔoul.dØ] that start with an initial empty onset containing a root node filled by a glottal stop. However, a default glottal stop cannot form the head of a branching onset: there are no forms such as *[ʔjoul.dØ] in English. What this suggests is that an unspecified root node cannot license a specified intraconstituent dependent.

[19] On this issue, see Harris' (1997) account based on the inheritance of depleted autosegmental licensing potential by prosodic dependents from their heads.

Interestingly, as Charette (1990) shows, word-internal empty nuclei in French must be phonetically expressed in certain contexts. Specifically, when the empty nucleus is preceded by a branching onset (e.g., *vendredi* 'Friday') or by a coda-onset sequence (e.g., *porcelaine* 'porcelain'), a schwa always appears. Under my proposal, this means that the empty nucleus must contain a root node whenever the onset it licenses has its own licensing responsibilities.

In addition, languages tend to avoid strings of consecutive empty nuclei that lack a root node and remain unexpressed. Instead, there is a decided preference for empty nuclei without a root node to be broken up by phonetically expressed nuclei (either filled nuclei or nuclei with a root node that harbour a default vowel). This may simply fall out, however, from a requirement that the head of a foot appear at regular intervals, so necessarily a filled nucleus appears to serve this function – Chapter IV addresses this issue in greater detail. Under this analysis, there is no need to appeal to proper government to determine the distribution of unexpressed empty nuclei; the pattern emerges as a consequence of foot structure.

Overall, there is a recurrent asymmetry in the distribution of filled and empty categories and, among empty categories, between those with and without a root node. Apparently, just as the distribution of filled segments is regulated by phonotactic well-formedness, the occurrence of empty categories with or without a root node is also constrained. The general tendency is for greater degeneracy to correlate with prosodic recessiveness and limited licensing responsibility. Unfortunately, more precise formulation of the phonological principles underlying the distribution of empty categories remains elusive. Whenever appropriate, however, this theme will be revisited in subsequent chapters, with an eye to further elucidating the constraints on the distribution of degenerate forms.

1.4 Conclusion

The proposed distinctions associated with the set of empty categories are highly abstract, but they serve to solve the problem of divergences in the behaviour of empty categories. These differences in behaviour may be observed both between languages and, more crucially, at times within a same language in identical contexts. As emphasized earlier, it is one thing to encounter different languages that treat certain phonological forms differently in similar contexts. Such a situation is at least potentially amenable to a solution based on parameter settings or constraint rankings. It is another matter entirely to encounter apparently identical phonological forms that behave differently within the same language. The idea that one parameter setting or constraint ranking might apply some of the time and another the rest of the time is highly implausible. An alternative but unappealingly ad hoc solution is to resort to lexical markings that isolate certain words from the regular phonological system. A preferable solution is to envisage a representational distinction that underlies the divergent behaviour. This is precisely the solution that is explored in the following chapters.

CHAPTER II

ELEMENT THEORY

2.1 Introduction

At first glance, this chapter may appear to digress from the main line of argument concerning empty syllabic categories, but the digression will prove both useful and necessary. It will be useful for subsequent chapters to have already established in detail how element theory works. Essentially, this chapter will develop a base for arguments later on. In particular, the element system will be important for determining:

i) the content of default segments that appear in empty categories;
ii) the nature of the segmental content that spreads to empty categories from adjacent segments;
iii) the complexity of segments (crucial for establishing whether sequences of consonants are well-formed branching onsets or coda-onsets).

For those less familiar with elements, which diverge in various ways from distinctive features, the chapter will serve as a valuable introduction. For those more conversant with element theory, which comes in different variants, this chapter will serve the necessary purpose of laying my element cards on the table, so to speak. I will propose some views (especially concerning the element @) which go counter to more standard positions.

Elements rather than distinctive features constitute the segmental primitives of Government Phonology. In the following sections, some emphasis will be placed on how elements differ from distinctive features, particularly when this has an impact on how elements interact with empty categories. For example, I argue that the elements @ and ? behave as universal default melody in nuclear and non-nuclear positions. These are the elements that (in my view) trigger either schwa or glottal stop in an empty root node. Clearly, the discussion concerning @ and ? will prove particularly pertinent to the main focus of this thesis.

Elements differ from the distinctive features of standard phonology in a number of ways. First, GP-style elements are strictly *privative*. Consequently, they are either present or absent in a segment. In accounts of distinctive features in the SPE-tradition, on the other hand, these figure constantly with either a plus or a minus value (nonetheless, cf. Trubetzkoy 1939; Rice & Avery 1991). Next, with a few exceptions, elements are *individually interpretable*. That is, while features need to be combined to define a surface form, elements can be realized phonetically on their own (on this, see Harris & Lindsey 1995). Elements also show a greater *freedom of occurrence* than features. All vocalic elements can appear in both nuclear and non-nuclear positions. Likewise, some consonantal elements can occur in nuclei. Finally, unlike features, element configurations are *headed*: complex segments contain a dominant element serving as head and recessive elements serving as dependents. Heads being obligatory, a simplex segment contains only a head.

These properties of elements have important consequences which will be returned to later. First, however, the elements themselves will be introduced. The next sections present the vocalic and consonantal elements in turn, showing how these are thought to correspond, either individually or in combination, to segments or 'phonemes' as represented by IPA symbols. Of particular interest is the elemental representation of two characteristically default segments, schwa [ə] and glottal stop [ʔ].

2.2 Element inventory

The system of privative elements as opposed to distinctive features is not exclusive to GP. The origins of element theory can be traced to Dependency Phonology (Anderson & Jones 1974; Anderson & Ewen 1980, 1987; Durand 1986a, 1995; Ewen 1995). A similar system is also advocated in Particle Phonology (Schane 1984, 1995) and in van der Hulst (1989).

Since the early days of GP, element theory has evolved, shedding in particular the distinction between charmed, charmless and neutral elements proposed in Kaye, Lowenstamm and Vergnaud (1985). Currently, there is no absolute consensus on either the set of primes or how they combine in complex segments. Diverging viewpoints can be found in Kaye, Lowenstamm and Vergnaud (1990), Harris (1990, 1994), Harris and Lindsey (1995, 2000), Rennison (1990), Kaye (2000), and Ingleby and Brockhaus (2002). Recently, Backley (2011) proposed a particularly streamlined version. I will not itemize all the points of contention, since many of these issues are orthogonal to our concerns. Instead, the focus will be on the view put forward in works by Harris, which constitute probably the most thorough-going account of element theory available. Certain modifications will also be proposed, particularly concerning the vocalic element @. Ultimately, these will have ramifications for our understanding of empty categories and the elements that may associate to them.

2.2.1 Vocalic elements

The core vocalic elements are A-I-U, with a fourth more marginal one being @. The term 'vocalic' is a bit of a misnomer, however, since A-I-U-@ can appear in both nuclear and non-nuclear positions. Still, these four elements are more quintessentially vowel-like in their phonetic (articulatory and acoustic) correlates. Certainly, A-I-U-@

are all that is needed to specify the vowel inventory of a language.[20] In fact, A-I-U are frequently all that is needed; many inventories do not use the element @. This difference between the elements A-I-U and @ will be a recurring theme here.

Vowels can be simplex, containing a single element, or complex, composed of two or more elements. The set of simplex vowels is presented in Table 2.1.

Table 2.1 *Vowels with a single element*

A	[a]	@	[ə]
I	[i]		
U	[u]		

Schwa, corresponding in my view to the @-element, is separated off for the simple reason that it behaves differently from the vowels [a i u]. First, almost all of the world's languages contain vowels corresponding to the elements A, I or U on their own. Many languages, on the other hand, lack schwa. A-I-U, in other words, are omnipresent elements; the element @ may be present or absent in a system. Next, the three vowels corresponding to A-I-U on their own constitute a common minimal vowel inventory. Many languages restrict their inventory to precisely the three vowels [a i u]: Classical Arabic, Tamazight, Aleut, Greenlandic, Quechua, Aymara, Warlpiri, and others. Interestingly, as inventories expand beyond these three basic vowels, they do not necessarily include schwa. We will thus put the element @ to one side for the moment and concentrate on the elements A-I-U.

Disregarding the element @, the inventory in Table 2.2 exhausts the possibilities for two-member complex vowels (the head element is placed first and underlined).

[20] An exception is nasal vowels, but their status as genuinely separate from their oral counterparts is debatable. Usually, the presence of nasality on a vowel can be traced either to an adjacent nasal consonant (for French, see Paradis & Prunet 2000) or to a harmony process (see Piggott 1992, 2003a).

Table 2.2 *Complex vowels with two elements*

$\underline{\mathrm{I}}$A	[e]	$\underline{\mathrm{A}}$I	[æ]	$\underline{\mathrm{I}}$U	[y]
$\underline{\mathrm{U}}$A	[o]	$\underline{\mathrm{A}}$U	[ɒ]	$\underline{\mathrm{U}}$I	??

For reasons that remain obscure, languages show a preference for the two-member vowels [e, o], with A as dependent, over the other two-member vowels in Table 2.2. The dispreferred vowels either have A as head [æ, ɒ] or they lack A entirely [y].[21] The hypothetical pair $\underline{\mathrm{U}}$I is not identified in the literature.[22] As an example of the preferred status of $\underline{\mathrm{I}}$A and $\underline{\mathrm{U}}$A, languages with five-vowel inventories typically augment the simplex set with [e, o] rather than with the vowels in (2b) or, indeed, with schwa – see, for example, Spanish, Serbo-Croatian, Maori, Hausa, Basque, Zulu, Hawaiian, and Kaigani Haida. That is, languages with precisely five vowels usually expand the basic inventory by adding the vowels [e, o].

In terms of three-member vowel configurations, again disregarding the element @, only the I-headed vowel shown in Table 2.3 is identified in the literature. The vowel

[21] Kaye, Lowenstamm and Vergnaud (1985) propose that in some vowel systems the tiers for two elements may be fused, which rules out their combination in a complex segment. For example, if the tiers for I and U are fused, the vowel [y] cannot occur in the inventory. This constitutes, however, only a partial solution, since systems with $\underline{\mathrm{I}}$A [e] necessarily accord separate tiers to I and A, so we are no closer to establishing why [æ] is dispreferred. See, however, the explanation given in Kaye, Lowenstamm and Vergnaud involving the notion of charm. See also the discussion on the markedness of vowels in chapter 9 of SPE (Chomsky & Halle 1968).

[22] The issue of $\underline{\mathrm{I}}$U versus I$\underline{\mathrm{U}}$ is in fact addressed in Kaye, Lowenstamm and Vergnaud (1985). In their system, elements correspond to fully specified feature matrices with one articulatory feature having a marked value, namely [-high] for A, [-back] for I, and [+round] for U. When elements combine, the dependent (or 'operator') contributes only its 'hot' or marked feature. Because the feature matrices that they provide for I and U differ based only on their marked feature, combining I and U with I as head or with U as head results in the same feature specification (i.e, the matrix corresponding to [y]). The element A differs not only based on its hot feature, but also based on its being [+low]. This permits A$\underline{\mathrm{I}}$ and $\underline{\mathrm{A}}$I to define distinct segments. No element has [low] as its hot feature, so this feature has a special status. Arguably, the inclusion of [+/-low] in the feature matrices undermines the very contention that elements are the primitives in phonological representations. It is tantamount to giving up on the whole project, since ultimately the feature [low] is the distinctive unit. (See the original for further details.) In Harris' (1994) system, on the other hand, elements are conceived of as cognitive categories with acoustic correlates (e.g., particular spectral peaks). (Again, see the original for details.) In this case, combining I and U should result in distinct output depending on which is head. Harris does not address this point explicitly, however.

[ø] appears for instance in French *feu* 'fire' and German *schön* 'beautiful'. Presumably the unidentified combinations in Table 2.3 are possible, but their identity has not been specified in the GP literature.

Table 2.3 *Complex vowels with three elements*

IUA	[ø]	UIA	??
		AIU	??

To recap, minimally, the set of vowels in Table 2.4 can be specified with just the elements A-I-U.

Table 2.4 *Vowels specified by* A-I-U

[i]	[y]			[u]
[e]	[ø]			[o]
	[æ]		[a]	[ɒ]

Now we can revisit two properties touched on earlier that distinguish elements from features: privativity and headedness. One of the advantages of elements being privative is that a distinction can be made between simplex and complex segments. With elements, there is a clear logic behind certain vowels (simplex) being preferred over others (complex) in vowel inventories.

In addition, as Harris (1994) emphasizes, neutralization of vowel contrasts in weak positions can be better captured with elements than with features. In element theory, neutralization involves reduction of vowels to the simplex set through element suppression. The reduction process typically occurs in unstressed syllables, and element suppression follows from the weak autosegmental licensing power of unstressed positions (Harris 1997). To illustrate, in unstressed nuclei in Bulgarian (Harris 1994) and Luiseño (Munro & Benson 1973), the contrasts between [e] and [i] and between [o] and [u] are neutralized such that only simplex [i] and [u] appear. In feature theory, such phenomena are portrayed as a change in feature value: [-high] → [+high]. Nothing in the feature system predicts that this should be the process that takes place, rather than the reverse: [+high] → [-high] or [i, u] → [e, o]. In element theory, such processes involve straightforward element suppression. When [e, o] → [i, u], complex IA and UA are reduced to simplex I and U. In both cases, the element A is removed from the expression. The reverse process of [i, u] → [e, o], involving unmotivated A insertion, is unlikely (if not impossible) in element theory. Viewed as feature change, however, both processes are equally plausible and equally arbitrary.

Interestingly, Belarusian (Barnes 2006) also has a reduction process targeting [e] and [o] in unstressed syllables. Here, however, the contrast between [e], [o] and [a] is neutralized, with all three being realized as [a]. That is, complex UA and IA are both reduced to simplex A. The process is identical in kind to what occurs in Bulgarian and Luseño, except that different elements are suppressed (U and I rather than A). In element theory, then, the two patterns are clearly related. In feature theory, however, the two processes involve arbitrary feature change. In addition, the feature changes involved in the Belarusian process are quite complicated: when [e] → [a], [-back, -low] → [+back, +low]; when [o] → [a], [+round, -low] → [-round, +low]. In brief, element theory is more effective in revealing how and why such phenomena occur.

A further advantage of privativity is that it reduces the number of possible phonological processes, making fewer predictions. Thus, while a system of bivalent features predicts that processes referring to, for example, both [+round] and [-round] should be possible, element theory predicts that only U (corresponding to [+round]) should be active. This is precisely what we find in harmony systems: the spreading of [+round]/U throughout a domain is found for example in Turkish (Clements & Sezer 1982; Charette & Göksel 1996), but I know of no language with [-round] harmony.[23] Similarly, languages such as Turkish with apparent [+/-back] harmony can be

[23] See Chapter IV for a more detailed discussion of harmony in Turkish from an element perspective.

reanalyzed to instantiate only I-harmony (corresponding to [-back]): when the element I is present in a domain, it spreads, and all vowels are fronted; in the absence of I, nothing occurs, and vowels remain back. The problem of bivalent primes overgenerating predicted processes has in fact been recognized in feature theory, leading to proposals that a number of features are monovalent (see e.g. Clements 1985b; Rice & Avery 1991; Lombardi 1995a, 1996; Pulleyblank 1995). Being inherently parsimonious, element theory avoids the issue of overgeneration entirely.

Element configurations are also headed. This property is important because it allows for expansion of segment inventories without adding to the set of primes. For example, by switching the head on a pair of elements, say changing $\underline{\mathrm{I}}\mathrm{A}$ to $\underline{\mathrm{A}}\mathrm{I}$, two separate segments can be designated: [e] and [æ]. Without this property, while [e] might be IA, [æ] would require a separate prime, say XA. Arguably, headedness also permits one element of a pair to be phonetically dominant: thus the relatively greater roundness of [o] in comparison to [ɒ] can be attributed to U being the head element in [o] ($\underline{\mathrm{U}}\mathrm{A}$) and the dependent element in [ɒ] ($\underline{\mathrm{A}}\mathrm{U}$). Finally, by invoking headedness, element theory attributes to segments the same asymmetrical structure thought to operate elsewhere in phonological representations. Branching onsets and rhymes, for example, are composed of a head and a dependent. Syllables and feet may be either strong (heads) or weak (dependents) (Liberman & Prince 1977; Halle & Vergnaud 1987). In element theory, subsegmental structure thus obeys the same organizational principle as prosodic structure (not to mention morphological and syntactic structure).

Earlier on, the vocalic element @ was put aside to focus on the elements A-I-U. Many languages construct their inventory using only the elements A-I-U. Commonly, then, a vowel inventory draws entirely on the vowels presented previously in Table 2.4. In other cases, however, the element @ is required, either on its own in the form of schwa or in combination with A-I-U.[24] This fourth vocalic element behaves differently from A-I-U in ways that will be explored in the next section.

2.2.2 The element @

As emphasized in Harris (1994), the element @, which replaces the cold vowel '*v*' of Kaye, Lowenstamm and Vergnaud (1985), can be thought of as expressing vocalic neutrality. In my view, on its own the element @ corresponds to schwa. Many phonologists working in the GP framework would probably not agree with this view: an alternative candidate for the element @ on its own is [ɨ]. Certainly, in Kaye, Lowenstamm and Vergnaud (1985), the 'cold' vowel is identified as representing a feature matrix with no marked features. This matrix defines the vowel [ɨ] (20).

(20) *Feature matrix for the 'cold' vowel 'v'* = [ɨ]

$$\begin{bmatrix} \text{-round} \\ \text{+back} \\ \text{+high} \\ \text{-ATR} \\ \text{-low} \end{bmatrix}$$

Schwa on the other hand corresponds to a combination of '*v*' and A, with the cold vowel as head and the element A as operator, which contributes only its marked (underlined) feature to the final expression (21).

[24] As will be discussed shortly, in my view, the element @ can be either present or absent in a representation and can contribute either as a head or dependent in a vowel. This contrasts with the view espoused elsewhere (notably in Kaye, Lowenstamm & Vergnaud 1985 and Harris 1994) that @ is omnipresent in expressions, appearing on any tier that is not used for a particular vowel. That is, with the vowel [a], @ would be latent on the empty I and U tiers. Under this view, @ contributes to the quality of an expression only when it has head status.

(21) *Feature matrix for schwa:* '*v*' + A = [ə]

$$\begin{array}{ccccc} \text{`}v\text{'} & & \text{A} & & [\text{ə}] \\ \begin{bmatrix} \text{-round} \\ \text{+back} \\ \text{+high} \\ \text{-ATR} \\ \text{-low} \end{bmatrix} & + & \begin{bmatrix} \text{-round} \\ \text{+back} \\ \underline{\text{-high}} \\ \text{-ATR} \\ \text{+low} \end{bmatrix} & = & \begin{bmatrix} \text{-round} \\ \text{+back} \\ \text{-high} \\ \text{-ATR} \\ \text{-low} \end{bmatrix} \end{array}$$

These identifications are largely maintained by Harris (1994) for the element @, (i.e, @ = [ɨ] and @A = [ə]). However, he is not entirely consistent in the phoneme-symbol/element-pattern correspondences he provides. To demonstrate, first Harris draws attention to the slightly different pronunciation of the unstressed second vowel in *Ros***a***'s* versus that in *ros***e***s.* He identifies the former unstressed vowel as [ə] or A@ and the latter as [ɨ] or @ (see p. 110).[25]

Immediately thereafter in the text, however, Harris discusses two types of reduced-vowel systems in English, the one (System A) which always uses schwa as the reduced vowel, and the other (System B) which uses schwa or, in certain contexts, [ɪ].

(22) *Two types of reduced-vowel systems in English*

a. Reduced vowel = [ə] (Systems A and B)

[ɒ]/[ɑ][26] *pr***o***duct-pr***o***duction, phot***o***graphy-phot***o***graph*

[ow] *ph***o***tograph-ph***o***tography*

[ɑ]/[æ] *photogr***a***ph-photogr***a***phy*

---

[25] In a stressed context, the combination A@ is said by Harris to underlie the vowels in words like *cut* and *third.* This analysis seems unlikely since, in non-rhotic dialects, the vowels in words like *third* and *thud* or *bird* and *bud* are distinct (i.e., the forms are not homophonous as would be expected). Indeed, this distinction is typically recorded in a pronouncing dictionary (e.g., Jones 1956) as: [θə:d] vs. [θʌd], [bə:d] vs. [bʌd].

[26] Where two possibilities for the unreduced stressed vowel are given here, the first corresponds to standard British, the second to standard North American English.

b. Reduced vowel = [ə] (System A), [ɪ] (System B)

[iː] *demon-demonic*

[aj] *horizon-horizontal*

[ɛ] *telepathic-telepathy*

What unifies the context in which System B selects [ɪ] as the reduced vowel is the presence of the element I in the unreduced counterpart. Vowel reduction in this instance can be captured (in Harris' view) as loss of all elements except I and/or promotion of latent @ to head status if necessary. Thus, when [ɛ] or @AI alternates with [ɪ], for example, the process involves just loss of A; when [iː] alternates with [ɪ], @ is promoted to head status. When a vowel alternates with [ə], on the other hand, nothing unifies the set of vowels that can be the unreduced counterpart. This is explained by treating the process as loss of all elements except @, which gains head status. The problem is that, in Harris' system, this should yield [ɨ], not [ə], since [ɨ] is @ and [ə] is @A. Of course, if [ə] is @A, this combination cannot be derived via straightforward element loss from the set of non-reduced vowel counterparts, since many of these lack the element A. If, on the other hand, [ə] is @, the same problem does not arise. Contradictions such as these have encouraged me to rethink the role of @ in the representations of vowels. Specifically, I propose that [ə] corresponds to @ on its own, whereas [ɨ] corresponds to @I.[27, 28] The rest of this section aims largely to establish the justifications for these claims.[29]

---

[27] Note that the potential solution of treating [ə] and [ɨ] as notational variants of the same vowel is not possible since various languages employ both, including Kpokolo (Kaye, Lowenstamm & Vergnaud 1985) and Chaha (Banksira 2000).

[28] The same analysis is adopted in Brockhaus (1995), that is, @ = [ə], although she stresses that the analysis is the subject of debate.

[29] Digressing somewhat, I think Harris is absolutely right in distinguishing between the unstressed vowels in *Rosa's* and *roses*, but in my dialect at least I associate [ə] with *roses*. The identity of the vowel in *Rosa's* is more difficult to determine. It may be an unstressed version of [ʌ]. Certainly, there is something special about absolute final unstressed vowels in English in that they tend not to be reduced to schwa. See for example *photo/shadow/minnow/yellow* and *pity/giddy/happy/money*. The flapping of [t] and [d] in the onset of the final syllable of these words is a clear sign of absence of

The element @ on its own corresponds to schwa in my view, and schwa is unusual in that it is not necessarily present in a vowel inventory. Hence not all languages employ the element @ on its own nor, as we will see, in combination with other elements. When schwa does appear in an inventory, it often has unusual properties (van Oostendorp 1995, 1999; Silverman 2011). Notably, schwa often alternates with zero: more than other vowels, schwa may be the target of deletion or insertion processes (Lombardi 2002b; Hall 2011). Typically, schwa is neither long nor stressed (though see Davis 1984 for stressable schwas in Squamish). In languages such as English, Russian (Crosswhite 2000), and Catalan (Wheeler 2006), unstressed vowels may reduce to schwa. Phonetically, schwa is a neutral vowel, with little to characterize it either in articulatory or acoustic terms.

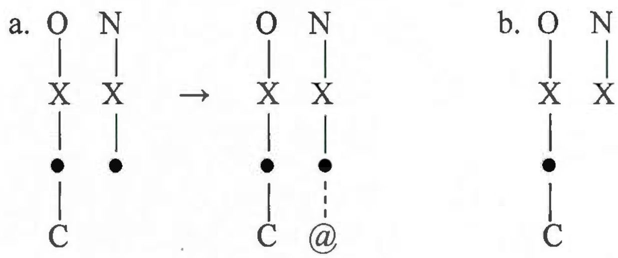
In my view, schwa is a typical default segment that is realized in an empty nucleus containing a root node. That is, while other elements must normally be specified in a representation in order to be present, @ can be inserted automatically into an unspecified root node.[30] The idea is that a root node, the unit of segmental structure that groups together all elements, can acquire @ without cost (23a). On the other hand, an empty nuclear position without a root node will not trigger @-insertion (23b). While the presence of a timing slot in the empty nucleus is sufficient to license the preceding onset, it fails to license default melody.

---

stress, yet the final vowels resist reduction to schwa (although they are all short, unlike their stressed counterparts). In brief, final unstressed vowels remain full in English. Hence it seems justified to associate the final vowel in *Rosa* (or *drama/saga/pita*) with a full vowel that resists reduction despite its unstressed status. Again, [ʌ] strikes me as a likely candidate. At this point, however, I am not sure what element content should be attributed to this vowel.

[30] Remember that in my view @ is not omnipresent in representations but is either absent or present (see previous note 24). Consequently, it can actually be inserted into a representation, rather than simply promoted to head status, as the more orthodox GP view would have it.

(23) *Schwa in an empty nucleus a) with a root node, but not b) without a root node*



The question of schwa will be returned to later on. For the moment we will concentrate on the role of @ in complex segments, whether as dependent or as head.

Clearly, the set of vowels whose element content is identified using A-I-U (the nine vowels listed previously in Table 2.4) is insufficient to account for all the vowels occurring in the languages of the world. For one thing, the IPA chart of vowels has twenty-eight members. Even taking into consideration the three unattributed element combinations given previously in Tables 2.2 and 2.3, a system limited to A-I-U is simply insufficient. Notably, the system fails to specify the lax vowels in Table 2.5.

Table 2.5 *Lax vowels*

[ɪ] [ʏ]	[ʊ]
[ɛ] [œ]	[ɔ]

The solution adopted here is to employ the element @: lax vowels have the same element content as their tense counterparts, but in addition they contain dependent @, as shown in Table 2.6.

Table 2.6 *The element content of tense and lax vowels*

Tense vowel	Elements	Lax vowel	Elements
[i]	$\underline{\mathrm{I}}$	[ɪ]	$\underline{\mathrm{I}}$@
[y]	$\underline{\mathrm{I}}$U	[ʏ]	$\underline{\mathrm{I}}$U@

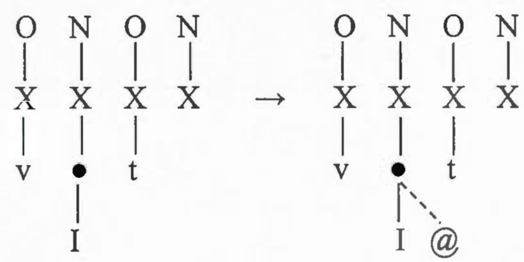
[e]	IA	[ɛ]	IA@
[ø]	IUA	[œ]	IUA@
[u]	U	[ʊ]	U@
[o]	UA	[ɔ]	UA@

Adding @ to the content of a segment is different from adding any of the elements A-I-U. For example, whatever their similarities, [i] and [e] feel like decidedly different segments. At least this is the case for me, an English speaker with distinctive [i] and [e]. On the other hand, whatever their differences, [i] and [ɪ] feel subjectively like quite similar segments, despite also being distinctive in English. In terms of the mental representation, the difference in these two vowel contrasts is that while [i] and [e] are distinguished by the element A, [i] and [ɪ] are distinguished by the element @. In essence, phonetically and cognitively, lax vowels with @ are attenuated versions of tense vowels. When @ is added to an expression, rather than contributing any positive characteristics, it dilutes the elements already there.

Various aspects of lax vowels are shared with schwa, a situation attributable to their having the element @ in common. For example, while schwa is the reduced vowel *par excellence* in English unstressed syllables, some dialects also employ [ɪ] as a reduced vowel (as we have seen) or even [ʊ] (Backley 2011). The choice between [ə], [ɪ] and [ʊ] as the reduced form depends on the quality of the full vowel. Note, for example, the possible alternations (due to stress shift) in the first vowels in **de***mon* [i] versus **de***monic* [ɪ] and **beau***ty* [u] versus **beau***tician* [ʊ]. Interestingly, in each case, the so-called reduced vowel is actually more complex than the full vowel: I → I@, and U → U@. The ability of @ to appear out of nowhere is also apparent when schwa forms the reduced vowel. See, for example, the alternation in the first vowel of **a***tom* and **a***tomic*: [æ] (AI) alternates with [ə] (@).

Schwa is thus a common epenthetic and reduction vowel; lax vowels are also common reduced forms. In other words, the element @ is at times added to a segment, with no local source, seemingly out of nowhere. Final vowel laxing in Quebec French provides a further example of this behaviour: the high vowels [i, y, u] typically surface as [ɪ, ʏ, ʊ] before final consonants. Thus, *vite* 'fast', *puce* 'flea', and *coûte* '(it) costs' are pronounced [vɪt], [pʏs] and [kʊt], whereas *vitesse* 'speed', *pucelle* 'virgin', and *coûter* 'to cost' are pronounced [vitɛs], [pysɛl] and [kute]. The derivation of [vɪt] (I@) from underlying [vit] (I) is illustrated below.

(24) *Derivation of* [vɪt] *from* [vit]



In my version of element theory, the process of laxing parallels that of schwa production in an empty root node shown previously in (23). In both cases, the element @ is inserted into the representation from nowhere. I know of no processes, on the other hand, where the elements A-I-U are added to underlying vowels without a local source.[31] As indicated for Bulgarian and Belarusian in the previous section, processes where they are removed from a segment in a weak position are relatively common, but not where they are added. The ability to appear seemingly out of thin air is thus a property especially associated with the element @.

[31] This does not preclude vowels other than schwa from appearing in epenthesis sites (Lombardi 2002b, Hall 2011). However, full vowels are usually employed only when a 'defective' vowel such as schwa or [ɨ] is unavailable. For example, Brazilian Portuguese, which disallows most consonants from word-final position, repairs loanwords such as *chic* [šik] by inserting a final [i]: [šiki] (Major 1986). Neither [ə] nor [ɨ] is found in Brazilian Portuguese. The same [i]-insertion phenomenon is found in loanwords to Cantonese, which also lacks [ə] and [ɨ] (Yip 1993).

Another aspect lax vowels share with schwa is that they are typically short. Certainly, a key part of the tense/lax dichotomy in English, German, Dutch, and Latin is that tense vowels are long and lax vowels are short. Relatedly, Plains Cree vowel length distinctions such as [i:]/[i] show up as tense/lax distinctions [i]/[ɪ] in East Cree. There appears to be an affinity of tense with long and of lax with short. This affinity is not absolute: Yurok has both long and short schwa ([ə:] and [ə]). Likewise, both tense and lax vowels in Telugu can be long or short (Wilkinson 1974). Still, I know of no languages that require tense vowels to be short and lax vowels to be long.

The cumulative weight of these connections between schwa and lax vowels supports my contention that they have something in common: namely, the element @. This is not the consensus view, however. Ingleby and Brockhaus (2002) and Backley (2011), for example, treat lax vowels as unheaded configurations. Thus, while [i] or [e] are $\underline{\mathrm{I}}$ and $\underline{\mathrm{I}}$A (headed), [ɪ] and [ɛ] are represented in their systems as I and AI (unheaded). To represent [ə], Ingleby and Brockhaus invoke the (consonantal) element R (unheaded). Backley represents [ə] as an empty element slot: | |.

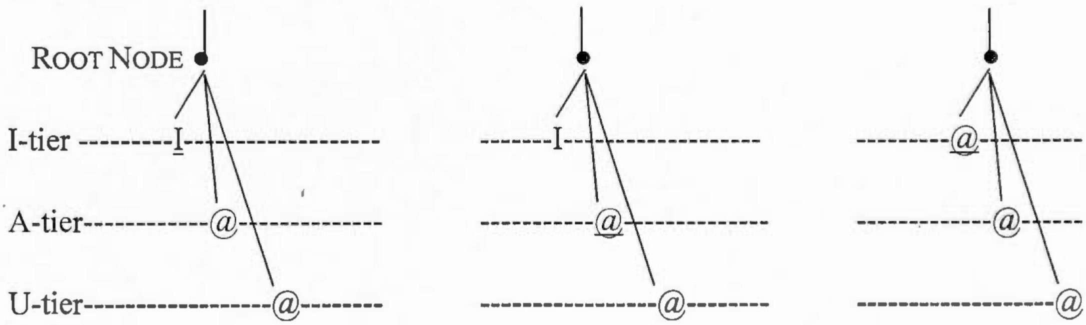
These views are problematic for two reasons. First, by employing unheaded segmental configurations, the parallel with the asymmetrical internal structure of branching constituents, syllables and feet is lost.[32] Next, by adopting different strategies to represent lax vowels and schwa, the link between the two is obscured: the similarity in their behaviour appears coincidental. The view that lax vowels and schwa have @ in common is thus clearly preferable.

A more promising view is expressed in Harris (1994) and Harris and Lindsey (1995). They adopt the position that the element @ is omnipresent in segments, but makes its

[32] The parallel between the asymmetry of constituents/syllables/feet and complex element configurations is in fact imperfect in that element configurations are not maximally binary (Heather Newell, personal communication). Hypothetically, segmental content may escape the maximal-binarity constraint on phonological licensing since, unlike constituents/syllables/feet, elements are realized simultaneously rather than sequentially such that locality/directionality is npot an issue.

presence felt only when it acts as head. More precisely, the special status of @ derives from its lacking its own autosegmental tier. Instead, @ is treated as latent on any unoccupied tier: for example, on the empty U and A tiers in the case of the vowel [i] – see (25a). (25b) has the same content as (25a), except that @ rather than I has head status, resulting in a lax vowel in this system. Schwa is represented as a segment containing the element @ as head and no other elements – see (25c).

(25) a. [i] b. [ɪ] c. [ə]



This view has certain advantages. For one, it unifies the representation of lax vowels and schwa. Consequently, it accounts for the dual processes of vowel laxing and reduction to schwa. Vowel laxing need not be viewed as insertion of @, but merely as promotion of latent @ to head position, a case of @ emerging from the shadows to contribute to vowel quality. Reduction to schwa can be viewed as element suppression with concomitant promotion of @ to head status.

The view also encounters important problems. First, because @ is deprived of its own tier, this restricts the total number of elements in a given segment to three: either AIU combined or else two of these elements plus @ as head. One crucial problem concerns the vowel [œ], occurring for instance in French *peur* 'fear' and German *Mörder* 'murderer'. This is the lax counterpart to [ø], a vowel containing the three-element configuration IAU. According to Harris and Lindsey's system, [œ] should thus employ the configuration @IAU. Yet @IAU is representationally impossible in this system, since @ is not accorded its own tier.

A more minor problem is that the system undergenerates the set of possible vowel configurations. Twenty-eight different vowels are included in the IPA chart, but only nineteen different combinations can be created with a system that insists on i) a maximum of three elements in a given form and ii) @ only contributing as a head. The combinations are shown in Table 2.7.

Table 2.7 *Possible element combinations in Harris and Lindsey's (1995) system*

Singleton elements:

A, I, U, @

Combinations of two elements:

AI, IA, AU, UA, IU, UI

@A, @I, @U

Combinations of three elements:

AIU, IAU, AIU

@AI, @AU, @IU

Either the IPA chart is too permissive or Harris and Lindsey's system is overly restrictive. One way of modifying the system is to add another element to the mix. A candidate for admission to the pantheon could be an ATR element. Such an element was included in the set originally proposed by Kaye, Lowenstamm and Vergnaud (1985), and Rennison (1990) offers some arguments in favour of ATR. For example, aside from permitting a wider inventory of vowels, inclusion of an ATR element permits us to capture examples of apparent +/-ATR harmony.

A clear problem (acknowledged in Kaye, Lowenstamm & Vergnaud 1985) is that tense vowels would then have a more complex composition than lax vowels. That is, tense vowels would be (literally) more marked, despite behaving as universally unmarked in vowel inventories. The logical solution is to invoke instead a –ATR element in representations (i.e., @, as we have been doing). This way, lax vowels are more marked than tense vowels. Also, cases of apparent +/–ATR harmony can be accounted for by having @ rather than ATR spread through the domain. When @ is present, all vowels become lax; when @ is absent, all vowels remain tense. Quebec French laxing harmony is a good candidate for such an analysis (Pòliquin 2006).

The solution I propose then is to accord @ its own tier and to allow it to contribute to an expression as both head and dependent. In this case, rather than figuring constantly as a latent element in segments, @ is either present or absent, just like any other element. By according @ its own tier, we make possible representations with all four elements, including the otherwise problematic vowel [œ] ($\underline{\text{I}}$UA@). By allowing @ to be either present or absent and to contribute to an expression as both head and dependent, we can also maintain the representation of tense/lax vowels shown previously and reproduced below in Table 2.8. Here the presence or absence of @ as dependent distinguishes between the tense and lax counterparts.

Table 2.8 *The element content of tense and lax vowel counterparts*

Tense vowel	Elements	Lax vowel	Elements
/i/	$\underline{\text{I}}$	/ɪ/	$\underline{\text{I}}$@
/y/	$\underline{\text{I}}$U	/ʏ/	$\underline{\text{I}}$U@
/e/	$\underline{\text{I}}$A	/ɛ/	$\underline{\text{I}}$A@
/ø/	$\underline{\text{I}}$UA	/œ/	$\underline{\text{I}}$UA@
/u/	$\underline{\text{U}}$	/ʊ/	$\underline{\text{U}}$@
/o/	$\underline{\text{U}}$A	/ɔ/	$\underline{\text{U}}$A@

The proposed solution also opens the door to distinctions based on @ as head or dependent. If @ operates as dependent in lax vowels, what vowels might be headed by @? A tentative proposal for two-member vowels is presented in Table 2.9.

Table 2.9 *@-headed combinations*

@I	[ɨ]
@U	[ɯ]
@A	[ɐ]

Some support for this being the set of @-headed vowels comes from cases of vowel epenthesis. Various epenthetic vowels are found cross-linguistically. We can exclude cases where epenthetic vowel quality is contextually determined through local spreading (e.g., in Selayarese or Khalkha Mongolian, as shown in (5) and (7) in the previous chapter). Instead, we focus on cases where a default vowel is inserted. The default vowel can be quite varied, but in general vowels from the proposed @-headed set are preferred to a full vowel. Lombardi (2002b) provides the following hierarchy for epenthetic vowels: ɨ > ə > i. In languages with both [ɨ] and [ə], [ɨ] will be preferred: for example, in Ethiopian Semitic languages such as Chaha (Banksira 2000) and numerous languages of Papua New Guinea (Foley 1986 – cited in Lombardi 2002b). Otherwise [ə] will be selected (e.g., in Dutch, Booij 1995, or in English), and then [i] (e.g., in Tongan, Kitto & de Lacy 1999, in Ponapean, Rehg & Sohl 1981, or in Brazilian Portuguese, Major 1986). Following my proposal for the element content of these segments, the hierarchy of default vowels is thus: @I > @ > I. Note that, under Harris' (1994) proposal, the element content for these vowels is: @ > @A > I. The relation between the preferred epenthetic vowels is thus less clear than in my system, which requires only @ and I singly or combined to capture the usual set. Clearly, other full vowels than [i] may be employed (e.g. [a] in Axininca Campa, Payne 1981, or [e] in Spanish), but the @-headed vowels [ɨ] and [ə], along with [i], are preferred.

Interestingly, Japanese lacks both [ɨ] and [ə], although it has [i] in its vowel inventory. Nonetheless [ɯ] (also in the Japanese inventory) appears by default in loanword adaptations: for example, 'cream' [krim] is realized as [kɯrimɯ].[33] If [ɯ] is similar to [ɨ] and [ə] in being @-headed, the selection of [ɯ] as default vowel in Japanese follows the pattern of favouring @-headed vowels in epenthesis sites. I know of no languages which select [ɐ] as default epenthetic vowel, however, so the evidence from epenthesis that [ə], [ɨ], [ɯ], and [ɐ] are all @-headed is incomplete.

At this point, a distinction should be made between two types of epenthetic vowel. The two types are found, for example, in loanword adaptations. On the one hand, there are vowels that are realized in empty nuclei with root nodes; on the other, there are full vowels that have been added to lexical representations. It is not always easy to distinguish the two, but I would argue that the schwa realized by English speakers between the first two consonants in *Gdansk* [gədænsk] is an empty nucleus with a root node that receives automatic realization as [ə]. The [i] that Brazilian Portuguese speakers realize at the end of *chic* [šiki] or the [ɯ] that Japanese speakers realize in *cream* [kɯrimɯ], on the other hand, correspond more probably to a lexically specified vowel.

Admittedly, the proposed system of four elements on four tiers, with the possibility of each element acting as head or dependent, generates more than the twenty-eight vowels found in the IPA chart. It may of course be that some combinations are simply unattested (i.e., we have yet to encounter a language that employs them). Other combinations may be illicit for reasons that are as yet unclear. Regardless, the system does an effective job of uniting schwa with lax vowels, thus accounting for their overlapping patterns of behaviour in reduction sites. It also links the @-headed set of vowels, [ə], [ɨ], [ɯ], and [ɐ], the first three of which are favoured as default epenthetic segments. These properties clearly count in its favour.

[33] Exceptionally, [o] is used in Japanese after [t] or [d], and [i] after [č] and [ǰ]. Otherwise, [ɯ] is the default epenthetic vowel.

One final issue needs to be addressed before we turn to consonantal elements. This concerns the low vowels [a] and [ɑ], which constitute something of an embarrassment of riches for GP. Many languages employ only one such low vowel, but confusingly transcriptions are not always consistent in which symbol is used for a given language. This does not mean, however, that [a] and [ɑ] are notational variants for the same segment, both simply corresponding to the element A on its own. For one thing, the two vowels are contrastive in some languages (see Quebec French *patte* [pat] 'paw' vs. *pâte* [pɑ:t] 'paste'). Consequently, there must be something to distinguish [a] and [ɑ]. I suggest that one contains the element A on its own and the other A as head with @ as dependent. Since *patte* and *pâte* are usually neutralized to [pat] in European French, and neutralization generally favours less marked output, [a] might be the better candidate for the element A on its own.

The notion that one member of the pair [a, ɑ] contains the element @ is certainly supported by examples from -ATR harmony. I attribute -ATR harmony to spreading of the element @. In some languages, -ATR harmony is triggered not only by the conventional set of non-low lax vowels but also by a low vowel. I suggest that a vowel with the elements A@ is the one that participates in harmony. For example, Tangale has both tense and lax vowels, transcribed as tense *i, u, e, o,* and lax *ị, ụ, ẹ, ọ, a* in Kiddo (1993). The sole low vowel patterns with the lax vowels in harmony, suggesting it has a -ATR element (i.e., @), which spreads from the stem – see the examples in (26).

(26) *Tangale laxing harmony*

tuk-ko	'hid'
ịl-ụ	'take (it)!'
sor-go	'your (m) height'
pẹd-nọ	'untie me'
pid-i	'the incisor'

sag-ọ	'ate'
kụlag-dọ	'her frying pan'

Interestingly, Akan has two low vowels, given as [ạ] and [a] by Clements (1985a). In vowel harmony, these pattern with the set of tense and lax vowels respectively (27).

(27) *Akan laxing harmony*

yarɪ	'to be sick'
kạri	'to weigh'
bisạ	'to ask'
pɪra	'to sweep'

In brief, the existence of a low vowel with the element A on its own, and another with A@, is supported by the evidence from Tangale and Akan vowel harmony.

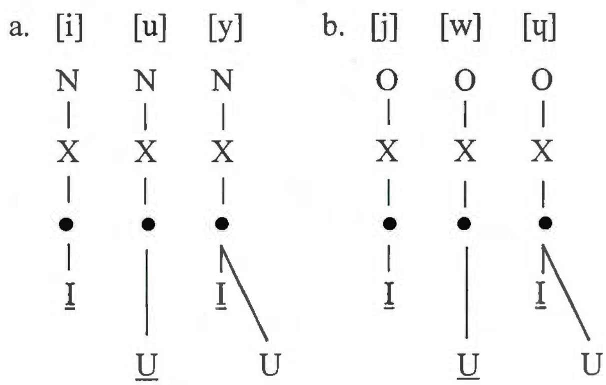
This concludes the section on vocalic elements. A number of modifications have been suggested for how the elements combine, with particular emphasis on the element @. In the view proposed here, @ must be accorded its own tier. It must also be allowed to figure in segments both as head (whether in schwa or in the complex vowels [ɨ], [ɯ], and [ɐ]) and as dependent in lax vowels. Crucially, @ is an element which can be inserted without a local source into a representation. This occurs, for example, in processes of vowel laxing. Finally, it is my contention that schwa, composed of @ on its own, is typically the realization of an empty root node under a timing slot associated to a nuclear position. Default insertion of segmental content such as @ is associated only with empty nuclei that contain a root node; empty nuclei with only a timing slot fail to trigger @-insertion. Hypothetically, this is because a root node is required in order to license default content.

2.2.3 Consonantal elements

Like 'vocalic', the term 'consonantal' is not entirely appropriate to describe elements. For one thing, the elements found in consonants sometimes crop up in vowels. Still, there are certain elements that are more commonly linked to non-nuclear positions. There is even less consensus about these non-nuclear elements than about vocalic elements, but again I am basing my account primarily on what is proposed in Harris (1990, 1994) and Harris and Lindsey (1995).

Before examining actual consonantal elements, however, we should consider the phenomenon of vocalic elements being realized as glides in non-nuclear positions. The vocalic and glide forms of I and U on their own and combined appear in (28).[34]

(28) *Vowels and their glide counterparts*

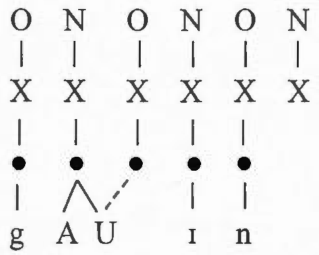


The notion that glides are the non-nuclear reflexes of vocalic elements is supported by the process of glide formation in hiatus: [j] or [w] emerge when the element I or U appears in one of a series of two vowels and spreads to the intervening empty onset. For example, a glide may be formed in the expressions *go in* or *see in.*

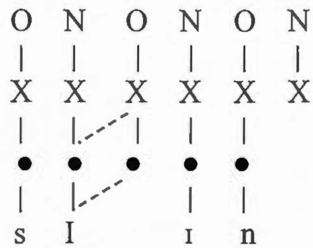
[34] Harris (1994) also proposes the velar approximant [ɤ] as the consonantal counterpart of @; the identity of A in a non-nuclear position is not indicated.

(29) *Glide formation*

a. *go* [w]*in*



b. *see* [j]*in*



Note that when a glide is formed via spreading of the element U from the complex vowel [o] (AU), the element U must be associated to two separate root nodes (as shown in 29a). Necessarily, then, the target contains its own root node. When a glide is formed from the simplex vowel [i], on the other hand, there are two possibilities (both shown in 29b): either the element I spreads to an empty root node in the adjacent onset or else the root node that supports I spreads to the timing slot in the onset. In other words, when the source is a simplex vowel, the target can be an empty onset either with or without a root node; when the source is a complex vowel, only a target empty onset with a root node can generate a glide. To my knowledge, this distinction has not been made previously in the literature, but it strikes me as crucial. The implications will be examined in greater detail in Chapter III.

Vocalic elements may occur also in combination with consonantal elements. In complex constructions, vocalic elements determine the place specifications of consonants. In addition, the element R, which on its own constitutes a tap [ɾ], is required to specify coronal place. Harris (1994) proposes the connections between elements and place presented in Table 2.10.

Table 2.10 *Element/Place correspondences*

Element	Place
U	Labial
R	Coronal
I	Palatal
@	Velar
A	Pharyngeal

The fact that vowels may assimilate to adjacent consonants and vice versa supports the view that the same elements are present in vowels and consonants. For example, roundness in vowels may be triggered by adjacent labial consonants in Turkish (Clements & Sezer 1982). In element theory, this involves spreading of the element U from the consonant to the vowel. Conversely, Japanese consonants are palatalized before the vowel [i] (Ito & Mester 1995), a process involving spreading of the element I from the vowel to the consonant.

In addition to the place elements, three manner elements and two laryngeal elements also contribute to consonant specifications. These are given in Table 2.11.

Table 2.11 *Manner elements and laryngeal elements*

Element	Manner property	Element	Laryngeal setting
?	Occlusion	H	Aspiration
h	Frication/noise	L	Voice
N	Nasality		

Individually, the element ? on its own corresponds to a glottal stop, and the element h to [h]. N on its own constitutes a placeless nasal, which cannot be realized as is. Typically, N appears on its own in coda position, receiving its place specification from an adjacent onset. H and L likewise are not individually realizable.

To give an idea of how the system works, a selection of element combinations and the segments they correspond to is provided in Table 2.12.[35]

Table 2.12 *Selection of segment specifications*

Segments:	[p]	[z]	[ŋ]	[r] (clear / dark)
Elements:	h, U, ?	h, R, L	N, @, ?	I, R / @, R

Using just these examples, a few patterns emerge. First, the element h is omnipresent in obstruents, but barred from sonorants. Next, the element ? occurs in both oral and nasal stops. Finally, the element L occurs in voiced obstruents, but not sonorants. The key element is ?, the special status of which is examined in the next section.

2.2.4 The element ?

Just as the vocalic element @ has special properties, so too does the consonantal element ?. Glottal stop, which is the realization of ? on its own, is the epenthetic consonant *par excellence* (Lombardi 2002a). Like schwa, however, glottal stop is commonly absent from phoneme inventories (i.e., it is not used contrastively). This includes English and German, even though glottal stop is used epenthetically in these languages. Unlike the element @, which may not be employed at all in a vowel inventory, the element ? is omnipresent in consonant inventories. This element adds occlusion to a segment, so it is present in stops. To my knowledge, all languages have stops, and stops are cross-linguistically preferred onsets. On the other hand, not all languages have fricatives (e.g., Warlpiri – Hale 1973), which lack the element ?.

In English, glottal stop production is fairly marginal: it occurs at the beginning of vowel-initial words, usually just in post-pausal position, as in [ʔ]*I see.* In German, on

[35] Although element configurations for consonants, as for vowels, are assumed to be headed, there is some uncertainty in the literature about which elements constitute the consonant head. Due to this incertitude, I do not indicate the head element here.

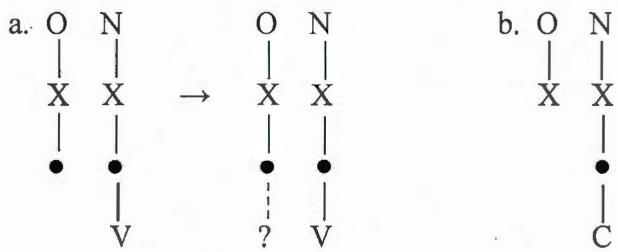
the other hand, glottal stops are produced in word-initial syllables and, word-internally, at the beginning of stressed syllables (Wiese 2000).

(30) *Glottal stop production in German*

Word-initial	Stressed syllable	Unstressed syllable
[ʔ]*Atem* 'breath'	*Po*[ʔ]*et* 'poet'	*Po_esie* 'poetry'
[ʔ]*offen* 'open'	*cha*[ʔ]*otisch* 'chaotic'	*Cha_os* 'chaos'
[ʔ]*Igel* 'hedgehog'	*The*[ʔ]*ater* 'theater'	*The_a* 'personal name'

It is my contention then that, just as @ appears automatically in an empty root node under a nuclear position, the element ? can be automatically generated in an empty root node under a non-nuclear position. This results in a glottal stop in German [ʔ]*Atem* or *Po*[ʔ]*et*. On the other hand, an empty non-nuclear position without a root node (e.g. in *Po_esie*) will not generate default melody, as shown below.

(31) *Glottal stop in an empty onset a) with a root node, but not b) without a root node*



Not all languages employ epenthetic glottal stops in empty root nodes, of course. As we saw earlier, these can also be filled by spreading of vocalic elements from an adjacent nucleus, resulting in glide production. In some cases, a full consonant can also be inserted into a representation where needed. A familiar example comes from Axininca Campa (Payne 1981), where [t] appears in hiatus position (e.g., i-N-koma-i → iŋkomat̲i 'he will paddle'). As Lombardi (2002a) emphasizes, however, when a consonant is actually inserted (i.e., not formed like glides via spreading of elements from an adjacent vowel), it is far more common for a glottal stop than [t] or any other consonant to appear.[36] This can be attributed to the special status of the consonantal element ?, which has the capacity to appear from nowhere in an empty non-nuclear position containing a root node.

2.3 Conclusion

To recap, unlike distinctive features, elements are entirely privative. This characteristic is invaluable in permitting us to calculate segment complexity and thus to determine the well-formedness of a pair of segments in a governing relation. Elements are individually interpretable, and they show great freedom of occurrence (e.g., vocalic elements in non-nuclear positions determine place or, on their own, they constitute glides). Although for the sake of simplicity the head of element configurations will not always be shown in subsequent chapters, the notion of headedness gives element theory a significant advantage: asymmetry is thus revealed to be a central organizing principle of both subsegmental and prosodic structures.

Vocalic elements are limited to A-I-U-@; consonantal elements to R-h-N-L-H-?. An important focus of this chapter was on the elements @ and ?. It was argued that these appear by default, sometimes in already filled segments (e.g., in the form of vowel laxing), but more importantly in empty nuclear and non-nuclear positions containing a root node. In the latter context, @ and ? are realized as schwa and glottal stop respectively. Empty categories with a root node are also potential targets of spreading of individual elements from adjacent segments. In the absence of a root node, in my view, a default segment cannot be inserted, but certain (more limited) forms of spreading, involving double association of a root node, are nonetheless possible. Finally, certain modifications were proposed concerning the element @: although @ has a primarily mitigative effect in complex segments, it is a bona fide element

[36] Paradis and Prunet (1991) indicate that coronal epenthesis, specifically of [t], occurs in Gokana and Amharic as well as in aphasic speech. The authors note nonetheless that glottal stops or glides are more common epenthetic consonants.

occupying its own tier (i.e., neither an empty element nor a tierless element). With these points established, we are now well-equipped to tackle in the coming chapters phenomena associated with two types of empty onsets, empty nuclei and empty codas.

CHAPTER III

EMPTY ONSETS

3.1 Introduction

From the GP standpoint, vowel-initial syllables start with an empty onset. An onset is required because onsets are heads, and heads are obligatory constituents. Empty onsets thus appear in two locations: in *initial* position of forms that start with a vowel (ØVCV); and in *medial* position between heterosyllabic vowels (CVØV). Not all languages allow onsets to be empty, but in those that do, we find evidence for two types of empty onsets of differing degrees of degeneracy. One piece of evidence involves 'phantom' consonants such as *h-aspiré* in French. The exceptional status of *h-aspiré* syllables has been abundantly addressed, both within the GP literature (notably Charette 1991) and beyond (e.g., Clements & Keyser 1983; Piggott & Singh 1985; Prunet 1986; Piggott 1991b; Gabriel & Meisenburg 2009). Charette proposes to distinguish between *h-aspiré* and vowel-initial syllables in French via the presence versus absence of a timing slot in the empty onset. I will demonstrate, however, that only a representation of empty onsets either with or without a root node is consistent with the behaviour of *h-aspiré* and vowel-initial syllables. These are illustrated below in Figure 3.1; for comparison, a filled onset containing specified content is also provided.

Empty onsets with a root node must be realized, essentially because the unspecified root node constitutes a vacuum that requires element content. For example, default melody such as a glottal stop may appear, in which case the element ? is automatically inserted into the receptive root node. This is the option shown in Figure

3.1. The empty onset without a root node lacks the requisite prosodic unit to generate default melody. Consequently, it often simply remains empty, again as illustrated.

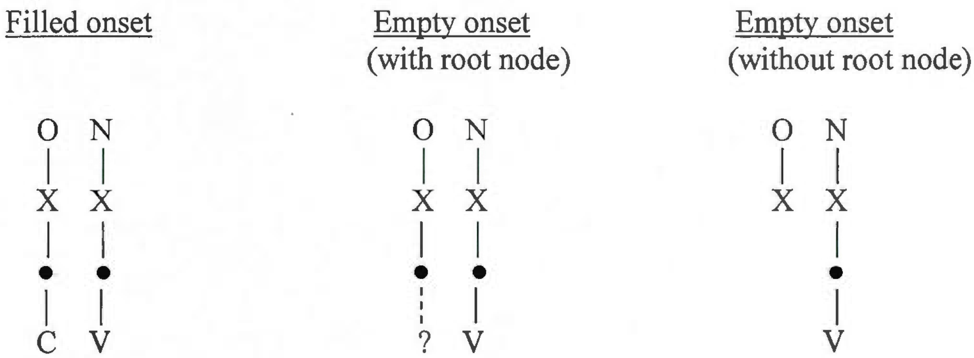


Figure 3.1 *Filled and empty onsets*

Another option that is available to empty onsets is for content to be supplied by spreading from an adjacent vowel. When empty onsets with a root node appear between heterosyllabic vowels (i.e., in hiatus), a glide [j] or [w] may be formed via spreading of the element I or U from one of the adjacent vowels. When glide formation involves the spreading of single elements from a complex source, the target necessarily contains its own root node to receive the select element, as shown below in Figure 3.2.

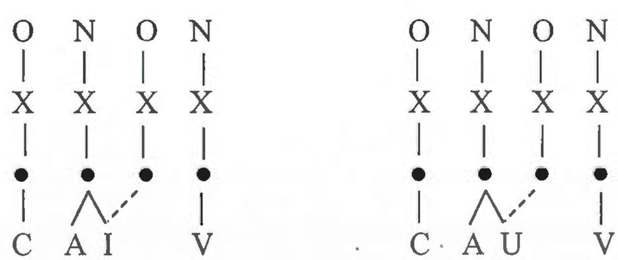


Figure 3.2 *Glide formation in empty onsets with a root node*

Glide formation is also possible when the empty onset lacks a root node, except that rather than it being the element itself that spreads, it is the root node under the adjacent nucleus that associates to the timing slot of the empty onset. In such cases, the source of spreading is restricted to simplex segments since it is the source root node itself that is doubly linked. This option is illustrated in Figure 3.3.

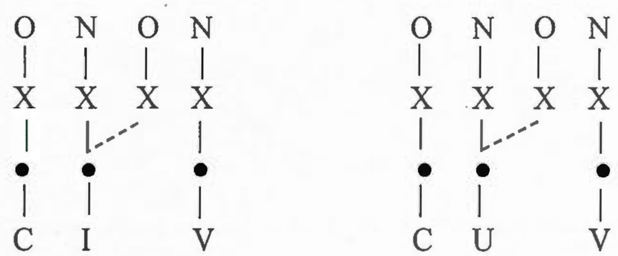


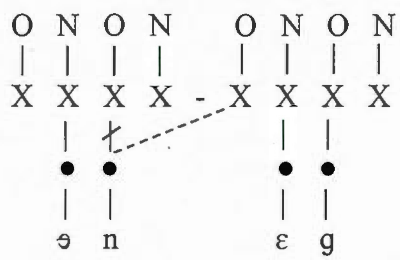
Figure 3.3 *Glide formation in empty onsets without a root node*

The existence of these two distinct types of glide formation constitutes powerful confirmation of the representational distinction I am proposing. Importantly, they cannot be accounted for with the representational distinction that has been proposed previously in the GP framework (i.e., involving presence/absence of a timing slot).

I assume that both types of glide formation constitute an option that languages are free to select or not. On the other hand, in the absence of glide formation, default consonant insertion in empty onsets with a root node is automatic. Consequently, onsets in hiatus remain empty only when i) the empty onset lacks a root node and ii) glide formation is not selected in the language (or else the context for glide formation is absent).

Another process that can affect initial empty onsets is the sandhi phenomenon of consonant re-association. This sometimes occurs at a boundary between words or morphemes when one form ends in a consonant and the subsequent form begins with a vowel. Under GP, the final consonant is of course analyzed as an onset of an empty nucleus, and, to complicate matters, an empty onset precedes the initial vowel. There are various ways of imagining the re-association of the final consonant such that it occupies the empty onset. I am not familiar with any cases where the consonant spreads such that a doubly linked structure is created, so this configuration will be ignored. Instead, I conceive of the process as involving some combination of spreading and delinking (or even erasure of one or more of the empty categories – see Harris 1994). I exclude the possibility that spreading of individual elements is involved, if only because this would require complex representations with multiple re-associations (at least, whenever the source consonant contains two or more elements). Instead, as plausible hypothesis, I adopt the view that the process involves delinking of the root node and associated elements from the first onset and re-association of this root node to the timing slot under the target empty onset (32).

(32) *Sandhi consonant re-association:* [ɘ.nØ-Øε.g] →[ɘ.nε.g] 'an egg'



The process of consonant re-association as represented in (32) would be favoured by empty onsets that lack a root node (as shown).[37] At the very least, the presence of an unspecified root node under the empty onset would be superfluous to consonant re-association; it could also conceivably block the process, given that the unspecified root node generates its own default content, thus obviating the consonant re-association process.[38]

---

[37] Vowel-initial words such as *egg* in English can of course at times be produced with a glottal stop, in which case the empty onset contains a root node. There are different ways of accounting for this variable behaviour. For example, we can posit a process of either fortition (insertion of a root node) or lenition (delinking of the root node) that applies in certain contexts. Otherwise, we can attribute the difference to two underlying forms, one with and the other without a root node in the empty onset. These two forms constitute different options that the speaker can draw on depending on such factors as phonological context and style. Subsequently, whenever this subject arises, I give preference to the fortition/lenition view, but nothing crucial hangs on this choice.

[38] The processes of default insertion, glide formation and consonant re-association do not exhaust the phenomena associated with vowel-initial syllables. In hiatus, other processes that are observed include:

Whether the empty onset before a vowel-initial syllable includes a root node or not may simply be a lexical choice in a language. In this case, the two types of empty onset can appear in the same context, even generating minimal pairs such as the *h-aspiré* word *haine* 'hate' and the vowel-initial word *aine* 'groin' in French. Both these words correspond to underlying [ɛn], but they behave differently, particularly with regard to various consonant re-association processes such as *liaison.* The question of *h-aspiré* syllables will be examined in detail in section 3.2. It will be argued that the distinctive behaviour of *h-aspiré* syllables suggests they start with an empty onset with a root node, whereas vowel-initial syllables start with an empty onset without a root node. To reiterate, this view is at odds with established GP representations, which propose a distinction based on the presence versus absence of a timing slot. It will be shown that this distinction is inadequate to account for the differences in behaviour of *h-aspiré* and vowel-initial syllables.

Another possibility is that a language may choose to use empty onsets with and without a root node in mutually exclusive contexts. This is what we find in German (section 3.3) and Malay (section 3.4). These languages show glottal stop insertion in one set of contexts, but in other contexts, they show either zero (in German) or glide formation via a doubly associated root node (in Malay). That is, depending on the context, the empty onset either does or does not contain a root node. To complete the picture, examples of glide formation via doubly associated individual elements (Shona and Japanese) are provided in section 3.5; in these cases, in contrast with Malay, the target of glide formation is an empty onset without a root node.

Interestingly, when the different types of empty onset appear in distinct contexts, the more robust empty onset endowed with a root node is favoured in prosodically or

---

deletion of one of the vowels; fusion or combination of the two vowels under a single nucleus via coalescence or diphthongization; and transformation of the first vowel into a glide – for examples and analysis, see Casali (1997), Rosenthall (1997), and Nevins and Chitoran (2008). Since these processes are not directly pertinent to the issue of distinguishing between empty onsets with or without a root node, however, they will not be addressed any further here.

morphologically prominent positions, namely at the beginning of stressed syllables or in root/word-initial position. The more degenerate empty onset lacking a root node tends to be relegated to recessive sites. From the perspective of licensing inheritance (Harris 1997), this distributional mismatch might be attributed to the depleted autosegmental licensing potential of non-head positions. Conversely, we might adopt the view that more degenerate forms are excluded from dominant sites, under a requirement that prominent empty onset positions contain a root node. Either way we formulate the explanation, a clear correlation emerges of prominent sites with less degenerate forms and of recessive sites with more degenerate forms.

3.2 French *h-aspiré*

In one and the same context, some languages show two types of vowel-initial syllables, which must be analyzed as preceded by two types of empty onset: those that are truly empty, and those that, while lacking specific content, behave much as though they were filled. This latter type of empty onset may contain what is often referred to as a phantom consonant.[39] This is not simply a consonant that alternates with zero, like [h] in the English words *pro<u>h</u>ibit* versus *pro(h)ibition.* Rather, phantom consonants are never present in the surface form, so there is no actual evidence for their identity. Consequently, learners cannot develop a precise underlying form for the consonant. Paradoxically, however, the behaviour of empty onsets with phantom consonants often resembles that of onsets with specified consonants. In brief, onsets with phantom consonants appear to occupy a middle ground between filled and truly empty onsets. This section explores the phantom *h-aspiré* consonant in French, and proposes that onsets with this elusive segment should be represented as containing an unspecified root node.

[39] For more on this topic, see the literature on phantom consonants in Aranese Gascon (Hualde 1992), Onandaga (Michelson 1986), Seri (Martlett & Stemberger 1983), Turkish (Clements & Keyser 1983), Maltese Arabic (Brame 1972), and Tiwi (Osborne 1974; Piggott 1991b).

French has two types of vowel-initial syllables: those that are truly vowel-initial, and those that are *h-aspiré*. While lacking an initial consonant, syllables that start with *h-aspiré* behave in many ways as though they were consonant-initial. In the majority of cases, *h-aspiré* syllables appear in word-initial position. Some examples contrasting vowel-initial words with *h-aspiré* words are given in (33). All of the examples are what may be thought of as homophonous minimal pairs. They are homophonous in that their underlying segmental content is identical. Nonetheless, when combined in phrases, however, *h-aspiré* words behave differently from true vowel-initial words.

(33)

Vowel-initial words		*H-aspiré* words		Underlying form
aine	'groin'	*haine*	'hate'	[ɛn]
eau	'water	*haut*	'high'	[o]
ache	'wild celery'	*hache*	'axe'	[aš]
une	'one'	*hune*	'top(mast)'	[yn]
ère	'era'	*hère*	'sorry fellow'	[ɛr]

All *h-aspiré* words began historically with a glottal fricative, which has been lost from the French inventory of consonants but which is preserved in the orthography. Some orthographically h-initial words are however treated as vowel-initial in French (e.g., *habit* 'clothes', *homme* 'man', and *hôte* 'host'). That is, all *h-aspiré* words begin with orthographic 'h', but not all words with orthographic 'h' are *h-aspiré*.

Vowel-initial words behave differently from consonant-initial words with respect to a number of processes, including *liaison, enchaînement,* selection of adjective or determiner form, and realization of a preceding empty nucleus. *Liaison* refers to the surfacing of latent word-final consonants, such as in the final [z] in *les* 'the' (pl) or the [t] in *petit* 'little' (masc). Such latent or 'floating' consonants are left unpronounced when the word is produced on its own or before a consonant-initial word as in *les bateaux* [lebato] 'the boats', but they surface before a vowel-initial word, as in *les aches* [le<u>z</u>aš]. Before an *h-aspiré* word, as in *les haches* [le_aš], latent consonants do not surface, just as before a consonant-initial word. The application of *liaison* before vowel-initial words has a straightforward explanation: the process supplies a consonant to an onset that would otherwise remain empty. The process fails to apply when a word starts with a consonant since the onset is already filled. For *liaison* purposes, then, *h-aspiré* words behave as though they too start with a filled onset, despite the absence of an underlying consonant.

When a word ends in a fixed consonant, this consonant can link onto a following word starting with a vowel but not with a consonant. Although this process of *enchaînement* does not typically occur before *h-aspiré* either, apparently it is not completely barred (this, according to Tranel 1995a, but cf. Boersma 2005). With respect to *enchaînement*, then, *h-aspiré* words usually behave as though they start with a filled onset, but at times as though they start with an empty onset.

The parallel between consonant-initial and *h-aspiré* words is maintained when it comes to selection of adjective or determiner form. Some masculine forms of prenominal adjectives have two allomorphs depending on whether the following noun begins with a consonant or a vowel. For example, in the case of *vieux/vieil* [vjø/vjej] 'old' (masc), the vowel-final allomorph is selected before consonant-initial nouns: *vieux chat* [vjøša] 'old cat'. The glide-final form appears before vowel-initial nouns as in *vieil ami* [vjejami] 'old friend', permitting the otherwise empty onset to be filled by the glide. Before *h-aspiré* words, the vowel-final rather than the glide-final allomorph is selected: *vieux hall* [vjø_al] 'old entrance hall'. Apparently, the glide is not required or cannot be incorporated into the *h-aspiré* onset.

When it comes to the form of the definite article *le/la* or *l'* (sing), again *h-aspiré* words pattern with consonant-initial words. Before consonants, *le* [lø] (masc) or *la* (fem) are used: *le chat* [løša] 'the cat'; *la pomme* [lapɔm] 'the apple'. Before vowel-initial words, on the other hand, *l'* [l] appears: *l'eau* [lo] 'the water' rather than *la eau* *[lao]. The use of *l'* instead of *la* before *eau* clearly permits the liquid to link up with the beginning of the noun and fill the otherwise empty onset. Before *h-aspiré* words, however, *l'* is not selected: *la hache* [la_aš], *le haut* [lø_o].[40] Once again the phonetic hiatus involving *h-aspiré* is tolerated, giving the impression that for the phonological system, the onset is somehow less empty than for vowel-initial words.

In addition, interestingly, the schwa in *le* and other clitics such as *ce* 'this' (masc) and *de* 'of' is obligatory before *h-aspiré*, whereas before a consonant it may be deleted. Hence: *pas d<u>e</u> hache* 'no axe', but *pas d(e) balle* 'no ball'. The same applies to word-internal schwas: in some contexts, they may be deleted before a consonant as in *en d(e)ssous* 'underneath', but not before *h-aspiré* as in *en d<u>e</u>hors* 'outside'.

There is some divergence in how empty nuclei behave at the end of words before *h-aspiré* and consonant-initial forms. In Standard European French, empty nuclei are often realized with a schwa-like vowel before *h-aspiré* words.[41] Before consonant-initial words, empty nuclei are usually realized only in specific contexts. Before vowel-initial words, empty nuclei are typically not realized at all. For example, the empty nucleus at the end of a determiner such as *cette* 'this' (fem) or a prenominal adjective such as *pauvre* 'poor' is frequently realized when followed by *h-aspiré*, more rarely when followed by a consonant, but never when followed by a vowel. The same behaviour is exhibited by the derivational prefix *re-*, which also contains an empty nucleus.

---

[40] I treat the use of *l'* as opposed to *le/la* as a question of allomorph selection, rather than of vowel deletion (the more common analysis). Nothing crucial hangs on this, although it permits a unified treatment of adjective and definite article forms. Specifically, by treating *l'* as a separate lexical entry that combines with vowel-initial words, I am analyzing it in the same way as adjective allomorphs such as *vieux/vieil* (i.e., [vjej] is not derived from [vjø]). Hence *le/la/l'* are analyzed like English *a/an* (*an* is not derived from *a* nor vice versa). The forms *le/la/l'* and *a/an* are suppletive. One justification for this view of *le/la/l'* is that vowels do not delete word-internally in French before other vowels: see *chaos* [ka.ɔs] *[kɔs] and *naïve* [na.iv] *[niv]. These forms contrast with **la ive* *[la.iv] 'the ground-ivy', realized as *l'ive* [liv]. In sum, a suppletion account is preferable to a vowel deletion account.

[41] The exact patterns may vary from one dialect of French to another (see Charette 1991 for the pattern involving Quebec French), but the same general tendency is found across varieties.

(34) *Final schwa production*

Before *h-aspiré*	Before a consonant	Before a vowel
cette haine [sɛtə̲ɛn] 'this hate'	*cette laine* [sɛt(ə)lɛn] 'this wool'	*cette aine* [sɛtɛn] 'this groin'
pauvre hère [povrə̲ɛr] 'poor sorry fellow'	*pauvre mère* [povr(ə)mɛr] 'poor mother'	*pauvre ère* [povrɛr] 'poor era'
rehausser [rə̲ose] 'to raise'	*retourner* [r(ə)turne] 'to return'	*rouvrir* [ruvrir] 'to re-open'

In brief, any analysis of *h-aspiré* onsets will have to account for their tendency to pattern with filled onsets in processes such as *liaison* and selection of adjective or determiner form. Nonetheless, it should ideally allow for the possibility of *enchaînement* before *h-aspiré* forms. It should also consider the divergent behaviour of empty nuclei before *h-aspiré* onsets versus filled onsets. For some reason, realization of a preceding empty nucleus is preferred before an *h-aspiré* onset.

A final issue concerns glottal stop production in *h-aspiré* and vowel-initial forms. Glottal stops may be produced variably at the beginning of both types of words. There is, however, a greater tendency for glottal stops to appear in *h-aspiré* words. Typically, a glottal stop is produced in a vowel-initial word only when the syllable receives special emphasis. For example, *incroyable* is normally pronounced [Øɛ̃krwɑ'jabl], with final stress; however, under particularly strong or emotional *accent d'insistance,* which involves stress shift to the initial syllable, a glottal stop typically appears in the empty onset, hence ['ʔɛ̃krwɑjabl] (Pagliano 2003). No such condition applies to *h-aspiré* words: a glottal stop may be produced regardless of stress status. This greater propensity of *h-aspiré* onsets to support a glottal stop is a further point that needs to be addressed.[42]

---

[42] The facts concerning glottal stop production in *h-aspiré* onsets are somewhat murky. Encrevé (1988) claims that a glottal stop never appears in combination with a preceding schwa (a position endorsed also by Côté 2007). Hence possible pronunciations of *une hausse* should include [ynə_os], with schwa,

A summary of the differences in behaviour between *h-aspiré*, consonant-initial and vowel-initial forms is provided in Table 3.1.

Table 3.1 *Behaviour of* h-aspiré, *consonant-initial and vowel-initial forms*

	H-aspiré	Consonant-initial	Vowel-initial
Liaison:	*No*	*No*	*Yes*
Enchaînement:	*No (variable)*	*No*	*Yes*
Adj, Det form:	*vieux, le/la*	*vieux, l(e)/la*	*vieil, l'*
Preceding schwa:	*Yes:* *cette haine* *rehausser*	*No (variable):* *cett(e) laine* *r(e)tourner*	*No:* *cettØ aine* *rouvrir*
Glottal stop:	*Yes (usual)*	*No*	*No (except under special emphasis)*

This is precisely the kind of situation that demands a representational distinction: two apparently identical forms that behave differently in the same context. Indeed, the notion of a representational distinction between the onsets of *h-aspiré* and vowel-initial words is well-established. In the linear tradition (Schane 1968, Selkirk 1972, Dell 1973), the behaviour of *h-aspiré* words was explained by positing an underlying consonant that fails to surface, either [h] or a glottal stop. In the nonlinear account of Piggott and Singh (1985 – see also Prunet 1986 and Gabriel & Meisenburg 2009), a similar solution was envisaged: *h-aspiré* words are treated as having an abstract underlying melody under the timing slot. This kind of explanation involving an underlying form that never surfaces constitutes absolute neutralization, thus posing a problem for learnability (Kiparsky 1968). The GP solution has been to represent *h-*

---

or [ynʔos], with glottal stop, but not *[ynəʔos], with both. However, Gabriel and Meisenburg (2009) report finding [ynəʔos] in their recordings of three out of twelve French speakers. Another three speakers showed only schwa production, and six showed only glottal stop production. Importantly, then, a full nine out twelve produced a glottal stop (with or without schwa), so the association between *h-aspiré* and glottal stop is quite strong. In these data, the association is in fact even stronger than that with schwa production (only six out of twelve speakers).

aspiré words with empty onsets with a timing slot and vowel-initial words with empty onsets without a timing slot (Charette 1991) – see Figure 3.4.

```
H-aspiré        Vowel-initial

  O N              O N
  | |                |
  X X                X
    |                |
    V                V
```

Figure 3.4 *GP representation of* h-aspiré *versus vowel-initial words*

Charette's distinction between *h-aspiré* and vowel-initial forms has certain advantages. Notably, it opens the door to an explanation for why the vowel in the definite article deletes before a vowel but not before *h-aspiré*: *la aine* → *l'aine* vs. *la haine*. The crucial factor according to Charette involves the adjacency of the nuclear timing slots in *la aine*. Given a sequence of two adjacent nuclear timing slots, an OCP violation occurs, and the first slot is hence deleted (35a).[43] When an *h-aspiré* empty onset intervenes, however, no such OCP violation incurs, since the intervening onset is itself endowed with a timing slot (35b).

(35) *Vowel deletion before vowel-initial but not* h-aspiré *forms* (l'aine *vs.* la haine)

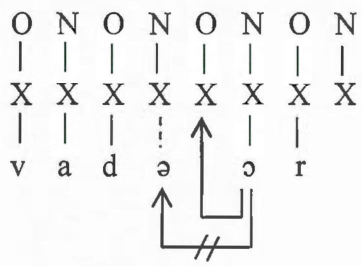
```
a. O N O N O N      O N O N      b. O N O N O N
   | |   | | |      | | | |         | | | | | |
   X X   X X X  →   X X X X         X X X X X X
   | ⍀   | |        | | |           | |   | |
   l a   ɛ n        l ɛ n           l a   ɛ n
```

In addition, Charette accounts for the requirement that the empty nucleus be realized in, for example, *va dehors* [vadəɔr] 'go outside' but not in *là-dessus* [ladØsy] 'on that' by arguing that the timing slot in the *h-aspiré* empty onset needs to be properly

[43] Unfortunately, Charette is silent on why this deletion process does not also apply to word-internal adjacent nuclei, as in *chaos* [kao] 'chaos', *Noël* [nɔɛl] 'Christmas' or *créer* [kree] 'to create'.

governed by the adjacent nucleus. In her view, a given nucleus can only properly govern one timing slot, so necessarily an empty nucleus preceding an *h-aspiré* empty onset will always be realized (36).

(36) *Proper government of an* h-aspiré *empty onset but not preceding empty nucleus*

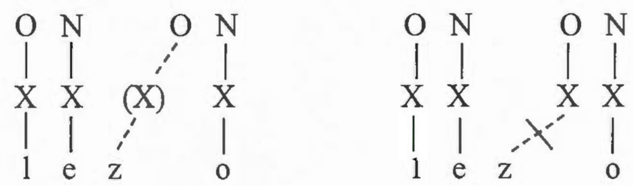


All the same, proper government cannot explain why the empty nucleus must be realized in *cette haine* [sɛtə̲ɛn] and *pauvre here* [povrə̲ɛr]. Being licensed by parameter, word-final empty nuclei do not require proper government, so the nature of the adjacent empty onset should be irrelevant to whether they are realized or not.

Likewise, Charette's representations encounter problems when combined with forms such as *les* [le(z)] or *petit* [pəti(t)] that end in so-called 'floating' consonants (i.e., latent segments that are not associated to any syllabic constituent). Remember that when the next word is vowel-initial, such floating consonants link up to the following empty onset, but when the next word starts with *h-aspiré*, the floating consonants fail to link up and thus remain unexpressed. Hence, *les eaux* 'the waters' is realized as [lezo], whereas *les hauts* 'the highs' is realized as [le_o]. In the GP tradition, floating consonants are represented as melody with neither a timing slot nor a syllabic constituent. When combined with the GP representation of *h-aspiré* versus vowel-initial words, however, this produces a troubling paradox: if vowel-initial forms lack a timing slot, but *h-aspiré* forms include one, vowel-initial forms (37a) should actually be less conducive to liaison than *h-aspiré* forms (37b).

(37) *GP representation of* liaison *with vowel-initial versus* h-aspiré *forms*

a. *les eaux* b. *les hauts*

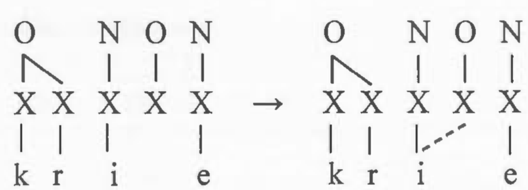


Under the GP set of representations, in order for *liaison* to occur before a vowel-initial form, a timing slot must be automatically generated between the floating final consonant and the empty onset (37a). However, in order for *liaison* to fail before an *h-aspiré* form, association of the floating consonant to the timing slot that is already in place inexplicably must be blocked (37b).

This paradox might be avoided by positing a timing slot above the liaison consonant itself. In this case, latent consonants would be floating with regard to the level of syllabic constituents rather than to the timing tier; the timing slot dominating the liaison consonant would spread to the vowel-initial empty onset, which lacks a timing slot, but not to the *h-aspiré* empty onset, which already contains one.

The problem is that a timing slot is independently required in the empty onset of vowel-initial syllables for the purposes of glide formation after high vowels: *crier* [krie] → [krije] 'to shout', *clouer* [klue] → [kluwe] 'to nail', *gluant* [glyɑ̃] → [glyɥɑ̃] 'slimy' (Kaye & Lowenstamm 1984).

(38) *Glide formation after high vowels in French:* [krie] → [krije]



Glide formation in these contexts involves spreading of the segmental content in the nucleus to the adjacent empty onset. That is, the segment is doubly associated to two timing slots, so necessarily the empty onset contains a receptive timing slot, as in (38).

In brief, it seems unavoidable that empty onsets in vowel-initial forms must contain a timing slot. The presence of this unit of structure is crucial to the processes of *liaison* and glide formation that target such forms. The next step then is to establish the representation of *h-aspiré* empty onsets. A possible representational distinction becomes available if we remind ourselves that the IPA symbols we have been using stand in for a more complex configuration involving a root node and associated element content. By introducing the root node as structural unit in representations, we can propose that empty onsets may be lexically specified either with or without a root node. *H-aspiré* words start with an empty onset containing a timing slot and root node; vowel-initial words start with an empty onset containing only a timing slot (Figure 3.5).

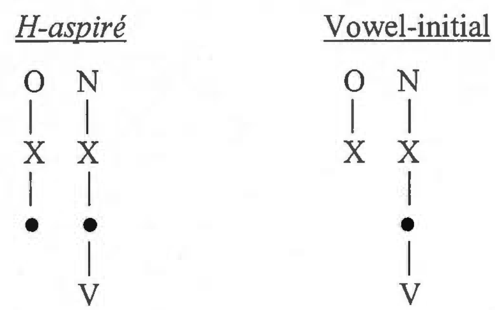
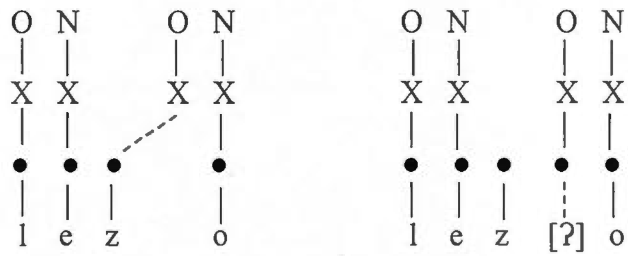


Figure 3.5 *Proposed representation of* h-aspiré *versus vowel-initial words*

Under this view, vowel-initial forms are perfectly suited to *liaison* (and glide formation) since they already contain the needed timing slot to harbour the root node of the floating consonant (39a). The presence of a root node in the *h-aspiré* form, on the other hand, blocks *liaison* and instead triggers glottal stop production (39b).[44]

(39) *Proposed representation of* liaison *with vowel-initial versus* h-aspiré *forms*

a. *les eaux* b. *les hauts*



Crucially, the proposed representations of two types of empty onsets with and without a root node are highly effective at accounting for *liaison,* whereas the traditional GP representations only generate a puzzling contradiction.

In similar fashion, the general blocking of the sandhi process of *enchaînement* can be attributed to presence of the root node. Likewise, before *h-aspiré* words, vowel-final allomorphs of the definite article *le/la* and of adjectives such as *vieux* are selected over their consonant-final forms *l'* and *vieil* since consonant reassociation is blocked when the empty onset contains a root node. Essentially, consonant reassociation involves spreading of the source root node to the target empty onset. This occurs only when the target lacks its own root node.

The presence of a root node in the empty onset also generally precludes the occurrence of a preceding unrealized empty nucleus. This is the case whether the empty nucleus appears in the same word (e.g., *dehors*) or at the end of a preceding word (e.g., *cette haine*). Charette proposes a viable explanation for the schwa in

[44] Presumably the root node also blocks glide formation. This remains an open question, however, as I am not aware of word-internal *h-aspiré* onsets that appear in glide formation contexts (i.e., preceded by a high vowel).

dehors based on proper government, but not for the schwa in *cette haine*. In my view, the production of schwa before *h-aspiré* conceivably points to a constraint on the adjacency of certain forms of empty category. Tentatively, schwa production is a means of avoiding the adjacency of an empty nucleus without a root node and an *h-aspiré* empty onset with a root node, whereas adjacency of an empty nucleus and empty onset both with a root node would be tolerated. Unfortunately, a more precise explanation of this apparent constraint remains elusive, and further research is needed to determine the validity of this tentative proposal.

More confidently, it can be proposed that processes of lenition or fortition apply variably to at times neutralize the contrast between the two types of onset. When the *h-aspiré* onset loses its root node, this removes the trigger for glottal stop production and opens the door for the linking phenomenon of *enchaînement.* For reasons that remain elusive, this lenition process does not however permit *liaison* to apply, nor does it affect selection of a determiner/adjective form. Arguably, the processes of *liaison* and selection of a determiner/adjective form could be ordered before the lenition process that removes the root node. Nonetheless, why the sandhi process of *liaison* but not *enchaînement* should be ordered before lenition remains to be seen.

Interestingly, under special emphasis, the empty onset in a vowel-initial word can undergo fortition, hence *incroyable* [ˈʔɛ̃krwɑjabl]. That is, a root node is added to the onset, resulting in glottal stop production. In this, French empty onsets exhibit behaviour similar to the pattern we will find in German in the next section: there is a correlation between empty onsets with a root node and stressed syllables. On the one hand, then, the representation of a vowel-initial word with an empty onset that does or does not contain an unspecified root node is simply a lexical choice in French. On the other hand, French at times shows an association between prosodically prominent sites and the less degenerate form of empty onset. This correlation between prosodic prominence and more robust segmental units is a common, even universal, pattern.

3.3 German glottal stops

In German (e.g., Wiese 2000), glottal stops do not appear in lexical representations. Unlike in Arabic, for example, glottal stops cannot be used contrastively; instead, they only occur as the default realization of an empty onset. Glottal stops appear at the beginning of vowel-initial roots and prefixes (40a). They even appear when these same vowel-initial roots and prefixes are themselves preceded by a prefix, regardless of whether this prefix ends in a vowel or a consonant (40b). Glottal stop insertion is thus pervasive in German: hiatus is not resolved by glide formation, and a preceding consonant does not re-associate to the empty onset.

(40) *Inserted root- and prefix-initial glottal stop in German*

a.	[ʔ]*árbeit*[45]	'work'
	[ʔ]*éin-flùss*	'influx'
	[ʔ]*ent-schéiden*	'to decide'
	[ʔ]*er-lésen*	'to select'
b.	*ge*-[ʔ]*àrbeitet*	'worked'
	be-[ʔ]*éinflùssen*	'to influence'
	[ʔ]*ún*-[ʔ]*entschìeden*	'undecided'
	[ʔ]*áus*-[ʔ]*erlèsen*	'choice, excellent'

There are, however, certain medial vowel-initial syllables in German that do not undergo glottal stop insertion. In root-internal hiatus, a glottal stop appears only if the vowel-initial syllable is stressed (41a).[46] Root-initial glottal stop insertion applies regardless of stress status (41b). (See also the previous examples in (40) of glottal stop insertion in both stressed (*éin-*, *ún-*) and unstressed (*ent-*, *er-*) prefixes.)

---

[45] Accents indicate primary and secondary stress.

[46] These are the patterns for standard German. Alber (2001) indicates, however, that in various southern varieties root-internal empty onsets never receive a glottal stop, regardless of stress status. In southern German, then, glottal stop insertion is restricted to root- and prefix-initial positions; all of the forms in (41a) would thus appear without a glottal stop in these varieties.

(41) *Root-internal and root-initial glottal stop insertion in German*

a. Root-internal (stressed/unstressed)		b. Root-initial (stressed/unstressed)	
Po[ʔ]*ét*	'poet'	[ʔ]*óper*	'opera'
cha[ʔ]*ótisch*	'chaotic'	[ʔ]*éwig*	'forever'
Klo[ʔ]*áke*	'sewer'	[ʔ]*áffe*	'monkey'
Po[Ø]*esíe*	'poetry'	[ʔ]*Attést*	'certificate'
Chá[Ø]*os*	'chaos'	[ʔ]*Idée*	'idea'
sáu[Ø]*er*	'sour'	[ʔ]*alléin*	'alone'

A similar pattern holds for vowel-initial suffixes: when these are unstressed (the usual pattern), they do not trigger glottal stop (42a). Instead, a stem-final consonant re-associates to the suffix onset. An exception is the secondarily stressed suffix *-àrtig*, which does trigger glottal stop production (42b).

(42) *Glottal stop insertion in vowel-initial suffixes*

a.	*kind-isch*	[kɪndɪš]	'childish'
	lebend-ig	[lebɛndɪç]	'alive'
	Sicher-ung	[zɪçərʊŋ]	'securing'
	Droh-ung	[droʊŋ]	'threat'
b.	*kind-àrtig*	[kɪntʔaɐtɪç]	'childlike'
	bös-àrtig	[bøsʔaɐtɪç]	'malevolent'
	einzig-àrtig	[ajntsɪçʔaɐtɪç]	'unique'
	neu-àrtig	[nɔjʔaɐtɪç]	'novel'

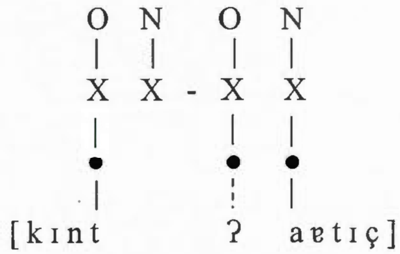
That the stem-final consonants in (42a) actually re-associate to the onset of the suffix is confirmed by the absence of obstruent devoicing. Normally, the onset of an empty nucleus devoices in German (Brockhaus 1995). Hence *Kind* 'child' is pronounced [kɪn̥tØ]; in *kindisch*, however, the voiced obstruent is able to surface since it occupies the onset of a realized nucleus. Similarly, word-final [ər] usually vocalizes to [ɐ].

Hence *sicher* 'sure' is pronounced [zɪçɐ]; in *Sicherung,* however, the combination surfaces as is, since the [r] is able to occupy the suffix-initial onset. Interestingly, even when there is no stem-final consonant to fill the empty onset of an unstressed suffix, no glottal stop is inserted – see *Drohung* [droʊŋ]. The hiatus and intervening empty onset are simply allowed to stand. In contrast, the stem-final consonants in (42b) above do not re-associate to the onset of the suffix; a glottal stop occupies the onset instead. This failure of the consonant to syllabify into the onset of the realized nucleus is confirmed by the application of final devoicing. Thus, in this case, the final obstruent in the root *Kind* undergoes devoicing in *kind-àrtig* [kɪnt̠ʔaɐtɪç]. Devoicing applies since the stop is in the onset of an empty nucleus. Likewise, *einzig-e* 'sole' (with unstressed suffix) is pronounced [ajntsɪgə], but *einzig-artig* surfaces as [ajntsɪçʔaɐtɪç]. Again, devoicing (and, in the case of a velar stop, spirantization) applies before the empty nucleus.

In my view, the differing behaviour of empty onsets in German points to a representational distinction: depending on the context, an empty onset either contains a root node or it does not. When the empty onset includes a root node, glottal stop production is automatically triggered, a consequence of it being the *raison d'être* of the root node to harbour melody. In the absence of a root node, glottal stop production cannot apply, but a preceding final consonant can re-associate to the empty timing slot. Glide formation fails to apply to empty onsets either with or without a root node; apparently, this option is not available in German. Empty onsets with a root node include those in root- and prefix-initial position, as well as in root-internal and suffix-initial stressed syllables. That is, they are associated with prosodically and morphologically prominent sites. Empty onsets without a root node are limited to root-internal and suffix-initial unstressed syllables (i.e., recessive sites).

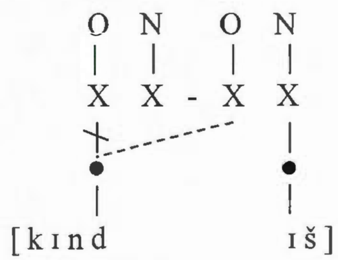
How this representational distinction works can be demonstrated with reference to stressed and unstressed suffixes. Remember that the former trigger glottal stop insertion, and hence should start with an empty onset with a root node. For example, underlying [kin**d-Ø**artıg], with the stressed suffix *–àrtig*, surfaces as [kin**tʔ**aɐtıç], with inserted glottal stop and devoiced onset of empty nucleus (43).

(43) *Empty onset with a root node before* -àrtig



Unstressed suffixes, on the other hand, fail to trigger glottal stop insertion. Instead, a stem-final consonant re-associates to the empty onset position, where it is able to maintain its voicing. This is shown for [kın**d-Ø**ıš], with the unstressed suffix *-isch,* which surfaces as [kın**d**ıš] (44).

(44) *Empty onset without a root node before* -isch



Note that consonant re-association in (44) necessarily involves spreading of the root node rather than of the elements themselves. The consonant [d] contains the elements h-R-?-L. In order to re-associate this content to an adjacent empty root node, four new association lines would need to be created. Under the assumption that phonological processes are preferably minimal, the account based on re-association of the root node to an empty onset lacking a root node is thus superior to one that invokes the spreading of elements. In brief, the evidence supports a distinction between empty onsets with and without a root node.

Interestingly, unlike in French, the different types of empty onset in German show different distributions.[47] Empty onsets with a root node are required in morpho-phonologically prominent sites, including in prefix-/root-initial sites and in stressed syllables; empty onsets without a root node only occur in unstressed suffixes and root-internal syllables. Indeed, related forms such as *Po*[ʔ]*ét*/*Po*[Ø]es*íe* and *cha*[ʔ]*ótisch*/*Chá*[Ø]*os* show actual alternation based on whether the empty onset occurs in a stressed or an unstressed syllable. The choice of empty onset is thus not arbitrary in German. This situation can be viewed from two angles. Hypothetically, word-internal unstressed syllables might be insufficiently strong in German to license an unspecified root node in an empty onset, an explanation which recalls Harris' (1997) notion of Licensing Inheritance. Conversely, it may be that the stressed syllable actually requires a more robust empty onset with a root node. In either case, there is a clear association between a recessive site and a more degenerate form, as well as conversely between a prominent site and a less degenerate form.

The portrait I have painted of the discrete behaviour of empty onsets in different contexts in German is, however, somewhat idealized. The reality is not so categorical as I have been indicating. Although frequently overlooked in the literature, strictly speaking, glottal stop production is variable at a prefix-stem boundary (45a), as well as at the beginning of words within an intonation group (45b) (Alber 2001). In both contexts, in the absence of a glottal stop, the preceding consonant apparently associates to the empty onset.

[47] Note that, because the differing behaviour of empty onsets in German occurs in mutually exclusive contexts, a representational distinction is not strictly required (although it is not excluded either). This is contrary to the situation in French, where only a representational distinction between two types of empty onset can account for the facts. My position is that when a particular representational solution is required in one language, it makes sense to extend this solution to other languages where a competing grammatical solution could be invoked. In essence, by adopting this approach we end up with a more parsimonious view of phonological systems.

(45) *Variable glottal stop production*

	With glottal stop	Without glottal stop	
a.	[ʔ]*án*-[ʔ]*orgànisch*	[ʔ]*án-orgànisch*	'inorganic'
	[ʔ]*áuf*-[ʔ]*erlègen*	[ʔ]*áuf-erlègen*	'to impose on'
	ver-[ʔ]*ántworten*	*ver-ántworten*	'to account for'
	[ʔ]*um*-[ʔ]*ármen*	[ʔ]*um-ármen*	'to embrace'
b.	[ʔ]*in* [ʔ]*Európa*	[ʔ]*in Európa*	'in Europe'
	der [ʔ]*Idée*	*der Idée*	'of the idea'
	[ʔ]*in* [ʔ]*Úngarn*	[ʔ]*in Úngarn*	'in Hungary'
	beim [ʔ]*Éssen*	*beim Éssen*	'while eating'

Interestingly, the realization without glottal stop is more common if the syllable containing the empty onset is unstressed ([ʔ]*án-orgànisch* or [ʔ]*in Európa*) than if it is stressed ([ʔ]*um-ármen* or [ʔ]*in Úngarn*) (based on data in Kohler 1994, cited in Alber 2001). In addition, unless they receive emphatic stress, post-verbal pronouns are almost never realized with glottal stop: *muss ich* 'must I', *habt ihr* 'have you' (pl), *kann er* 'can he' (Hall 1992). Presumably, these pronouns are cliticized and behave phonologically more like (unstressed) suffixes than like separate words. Similarly, grammatical words in general within an intonation group are more inclined to lack glottal stop. For example: [ʔ]*an* [Ø]*einem* [ʔ]*Ábend* [Ø]*in* [ʔ]*Úngarn* 'on an evening in Hungary' (Hall 1992).[48] In a nutshell, glottal stop production in a possible context is sensitive to how prosodically and morphologically prominent the particular site is.

Variable phenomena are of course problematic for any grammatical paradigm. The inadequacy of accounts attributing variation to style-dependent multiple grammars led to a proposal for variable rules with relative frequencies of application (Cedergren & Sankoff 1974). Within the principles and parameters framework, I am not familiar

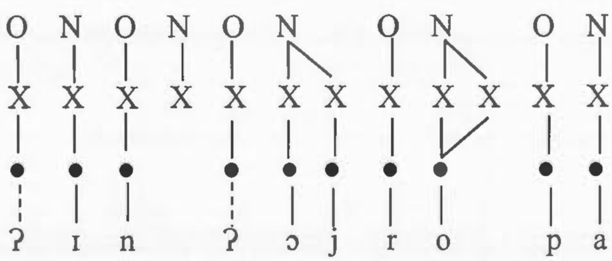
[48] Even though the final consonant in *Abend* may link up with the following vowel-initial word, final devoicing is obligatory all the same. Apparently, devoicing applies before re-association in this case.

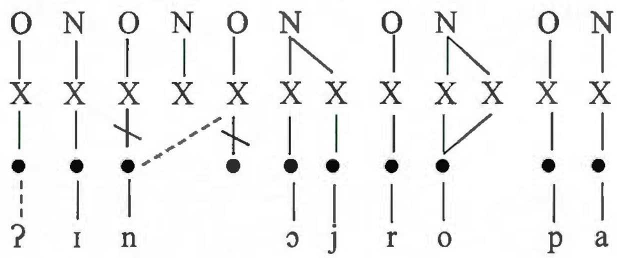
with any attempt to model variation. Necessarily, variation is unsettling for a paradigmatic approach since paradigms should be fixed for a language; they cannot fluctuate from one setting to another. Likewise, constraint-based grammars (as usually conceived) are inherently categorical: the constraint hierarchy always selects a single optimal candidate for output. Consequently, one of the criticisms leveled at OT has been that it is incapable of handling variation. A number of solutions have been proposed to the problem of encoding variation within an OT grammar: Kiparsky (1993) revised the multiple-grammars explanation from an OT angle; Reynolds (1994) and Anttila (1997) accounted for variation via the crucial non-ranking of constraints; and Boersma and Hayes (2001) proposed a stochastic version of OT involving a continuous ranking scale rather than a constraint hierarchy of strict domination.

From a representational perspective, variation can be captured as changes in the representation as a result of fortition and lenition processes that variably insert or delink prosodic structure. In German, it is not the case that contexts typically associated with absence of glottal stop sometimes receive a glottal stop. Consequently, fortition does not play a role here. It is the case, however, that contexts associated with glottal stop insertion sometimes fail to undergo the process. For example, in the phrase *in Europa,* either the empty onset before *Europa* is realized with a glottal stop or else the preceding consonant re-associates. This latter pattern can be attributed to a lenition process involving loss of the root node in the empty onset. The derivation of the two possible surface forms are shown in (46a,b).

(46) *Variable glottal stop insertion*

a.



b. 

In (46a), the two initial empty onsets with unspecified root nodes surface as expected with default glottal stops: [ʔ]*in* [ʔ]*Europa.* In this case, the ʔ-element is supplied by default. In (46b), on the other hand, the root node is delinked from the empty onset at the beginning of *Europa.* This removes the context for default glottal stop and triggers instead re-association of the preceding consonant into the empty onset. In brief, both the typical and variable patterns of glottal stop insertion and consonant re-association are amenable to a representational account based on different degrees of degeneracy in empty onsets. That these empty onsets may vary based specifically on the presence versus absence of a root node is also justified: the root node as prosodic unit constitutes a plausible trigger for default melody such as a glottal stop; conversely, an empty onset with a timing slot constitutes a suitable target for consonant re-association via spreading of an adjacent root node. Further evidence for this approach comes from phenomena in Malay, which instantiates both glottal stops and glides in empty onsets.

3.4 Malay glottal stops and glides

The evidence from Malay is not straightforward, but when teased apart, a picture emerges where positing two types of empty onsets helps account for various particularities in their behaviour. Malay (Durand 1987; McCarthy & Prince 1993; Tajul 2000; Zaharini 2005) employs two strategies for empty onsets in hiatus: glottal stop insertion and glide formation.[49] According to Zaharini (2005), both processes are limited to hiatus position. Based on Maris (1980), on the other hand, Durand (1987) indicates variable glottal stop insertion before vowel-initial forms. Thus, *apa* 'what' can apparently be realized as either [ʔapa] or [apa]; *itik* 'duck' as either [ʔiteʔ] or [iteʔ].[50] Otherwise, the literature shows general agreement as to the distribution of glottal stops (47) and glides (48).

In hiatus between a prefix and root, glottal stops appear to the exclusion of glides.

(47) *Glottal stop insertion in hiatus between prefix and root*

di-ukir	→	di[ʔ]ukir	'be carved'
di-ikat	→	di[ʔ]ikat	'be tied'
di-ankat	→	di[ʔ]ankat	'be lifted'
kə-enakan	→	kə[ʔ]enakan	'thrilled'
sə-elok	→	sə[ʔ]elok	'be as pretty as'
api-api	→	api[ʔ]api	'fires'
ǰuru-ačara	→	ǰuru[ʔ]ačara	'master of ceremonies'

Glottal stops appear in this context even when [i] or [u], which constitute potential input for [j] or [w] glide formation, appear as initial vowel. Note that in non-hiatus context, when the prefix ends in a consonant, no glottal stop appears before the root: tər-elok → [tərelo ʔ] 'most beautiful', pəŋ-aŋkat → [pəŋaŋkat] 'lifter'. In this case, the final consonant re-associates to the empty onset, removing the context for glottal stop production.

[49] Unlike in languages such as Arabic, there is no contrast between stable and inserted glottal stop in Malay. Wherever a glottal stop appears, then, it is best seen as the realization of an empty onset. An exception may be in word-final position, where [ʔ] is the surface realization of [k], which contains the elements h-@-ʔ. In this case, the glottal stop could arise from suppression of the elements h-@ in the velar stop, leaving only the element ʔ.

[50] Tajul (2000) certainly indicates a glottal stop before vowel-initial forms following a vowel-final form (hence, in hiatus at a word-boundary). For example: *hari* [ʔ]*ini* [ʔ]*ada kaliah* 'today there is a lecture'.

In hiatus between a root and suffix, however, the glides [j] or [w] are formed when [i] or [u] appear as the first vowel (48a). Otherwise, a glottal stop is inserted (48b).

(48) *Glide and glottal stop insertion in hiatus between root and suffix*

a.	uǰi-an	→	uǰi[j]an	'test'
	pər-kəlahi-an	→	pərkəlahi[j]an	'fight'
	bantu-an	→	bantu[w]an	'aid, relief'
	aku-i	→	aku[w]i	'acknowledge'
b.	mula-i	→	mula[ʔ]i	'begin'
	məŋ-gula-i	→	məŋgula[ʔ]i	'cause to sweeten'
	kə-raǰa-an	→	kəraǰa[ʔ]an	'government'

Again, when the root ends in a consonant, no glottal stop or glide appears before the vowel-initial suffix: milik-i → [miliki] 'cause to own'.

In brief, empty onsets at the beginning of roots and suffixes behave differently. Before vowel-initial roots, glottal stop insertion applies but not glide formation. I propose that these roots begin with an empty onset containing a root node, which triggers default insertion of the ʔ-element. Normally, this empty onset is realized with a default glottal stop as shown in (49) for *di*[ʔ]*ukir* (from underlying *di-ukir*).

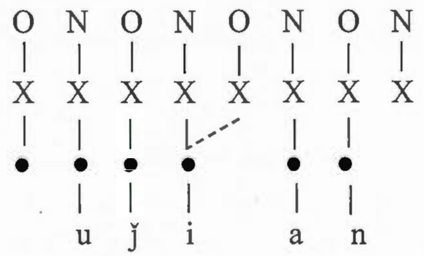
(49) *Root-initial empty onset with root node*

```
O  N  O  N  O  N  O  N
|  |  |  |  |  |  |  |
X  X  X  X  X  X  X  X
|  |  |  |  |  |  |
•  •  •  •  •  •  •
|  |  :  |  |  |  |
d  i  ʔ  u  k  i  r
```

An exception to glottal stop insertion is when the root is preceded by a consonant (or perhaps by a pause). In this case, the root node in the empty onset is apparently delinked, removing the context for insertion of the ?-element and thus for glottal stop production. This lenition process opens the door for the preceding consonant to re-associate to the initial onset position via a process of spreading of the root node.

Empty onsets at the beginning of suffixes are realized as glides whenever circumstances are favourable. Glides in Malay can only be formed after the vowels [i] or [u], never after [e] or [o].[51] The absence of glide formation after [e] and [o] suggests that the process must involve spreading of the root node of the previous vowel to an empty onset that itself lacks a root node.

(50) *Suffix-initial empty onset without root node:* uǰi[j]an



If glide formation were possible following [e] or [o], necessarily this would involve double association of the element I or U to two root nodes, not of the root node itself to two timing slots. If the empty onset before a suffix lacks a root node, this would motivate the limitation of glide formation to instances where [i] and [u] precede. When a consonant-final root precedes the suffix, the consonant itself occupies the empty onset via the same kind of association process behind glide formation: the root node dominating the consonant spreads to the timing slot under the empty onset.[52]

[51] The same restriction applies in French: glide formation applies after high vowels, as in *crier* [krije] 'to shout', *clouer* [kluwe] 'to nail' or *gluant* [glyɥɑ̃] 'slimy' (as mentioned previously in section 3.2); it does not apply after mid vowels, as in *créer* [kre_e] 'to create' or *boa* [bo_a] 'boa'.

[52] To complicate matters, Kassin (2000) in fact shows a consonant that links up to a suffix as becoming long (hence: lətop-an → [lətoppan]), but a consonant that links up to a root as remaining short. Why only the former should be able to occupy two skeletal slots and not the latter escapes me. In addition, Kassin refers to the process as gemination, but according to GP, final consonants are onsets of empty nuclei, so these long consonants would normally be a sequence of two onsets, separated by an empty

When the root ends in a vowel other than [i] or [u], the onset undergoes a fortition process. A root node is added to the empty onset and a glottal stop is produced: mula-i → mula[ʔ]i.

The representations for vowel-initial roots and suffixes are thus underlyingly discrete, which explains their divergent behaviour. Root-initial empty onsets contain root nodes in the lexical representation, whereas suffix-initial empty onsets lack a root node (essentially the same correlation as was found for German). Nonetheless, under certain circumstances, each type of empty onset can undergo a process of either lenition (loss of the root node) or fortition (acquisition of a root node). This explains the instances where the behaviour of the empty onsets in the two contexts overlaps.

Support for the view that root-initial empty onsets contain a root node comes from a language game (Tajul 2000; Zaharini 2005). In the version of the game given in Zaharini, the first two syllables of a root are inverted (51a). Interestingly, when the first syllable in a vowel-initial form is inverted, it is realized with a glottal stop (51b).

(51) *Malay syllable-inversion game*

a.	batu	→	tuba	'stone'
	satu	→	tusa	'one'
b.	apa	→	paʔa	'what'
	aku	→	kuʔa	'I'
	itu	→	kuʔi	'that'

Of particular interest is the presence of glottal stop in [kuʔa] and [kuʔi]. The output of the inversion process for [aku] and [itu] is [ku.a] and [ku.i], with a hiatus preceded

nucleus. Further complicating the situation is the fact that final [k] is realized as [ʔ]. Then, when [k] spreads and lengthens, it becomes a sequence of glottal stop + velar stop [ʔk]. Normally a doubly linked root node should not change its colours in midstream, as evidence for geminate integrity notably demonstrates. Still, these issues are orthogonal to my main concerns, so I will consign them to future research.

by a high vowel. This is precisely the context for glide formation between a root and suffix. We might thus expect the context to lead to the surface forms *[kuwa] and *[kuwi]. The fact that these forms are illicit, however, is consistent with an analysis of root-initial empty onsets as containing a root node. The presence of the root node acts as automatic trigger for a glottal stop.

Although less clear-cut than what we saw in German, the evidence from Malay is nonetheless suggestive of a need to distinguish between two degrees of degeneracy in empty onsets: those that contain a root node and trigger glottal stop production versus those that lack a root node and lead to glide formation via spreading of the root node under an adjacent vowel. Interestingly, some languages show another pattern of glide formation, in which the target is better conceived of as an empty onset with a root node. Evidence for this process can be found, notably, in Shona and Japanese.

3.5 Shona and Japanese glides

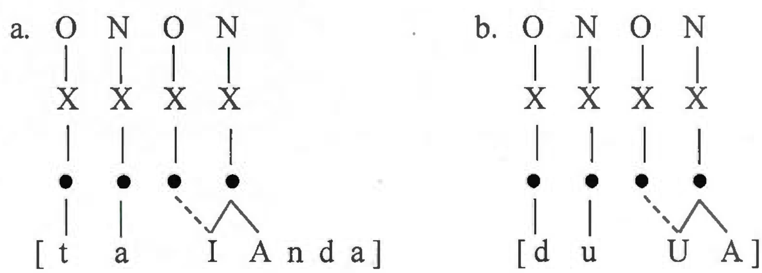
Not all languages restrict glide formation to empty onsets adjacent to [i] and [u]; some languages allow glide formation when the adjacent vowel is, for instance, [e] or [o]. Shona shows regressive formation of the glides [j] and [w] before [i]/[e] and [u]/[o] (Mudzingwa 2007, cited in Zygis 2010).

(52) *Regressive glide formation in Shona*

ta-end-a	→	ta[j]e^{n}da	'we went'
či-it-o	→	či[j]ito	'an act'
ta-i-p-a	→	ta[j]ipa	'we gave it'
duo	→	du[w]o	'fish trap'
a-or-a	→	a[w]ora	'they rot'
ama-uta	→	ama[w]uta	'deaf person'

Cases of glide formation involving [e] and [o] cannot be represented as spreading of the root node belonging to the adjacent nucleus, since the root node harbours the elements I-A and U-A respectively. Instead, only the elements I and U are doubly associated. That is, I or U is linked to the root nodes under both the nucleus and the empty onset. In the above forms in Shona, then, the empty onsets in hiatus presumably contain a root node. This root node is filled not by default insertion of the ?-element, but by spreading of the element I or U from the adjacent nucleus.

(53) *Shona glide formation through spreading of the element I or U*



Shona is not an isolated case. The same phenomenon of regressive glide formation in hiatus is found in Yucatec Maya: *le me:sa-e?* → *le me:sa*[j]*e?* 'that table'; *le me:sa-o?* → *le me:sa*[w]*o?* 'that table over there' (Orie & Bricker 2000).

Glide formation from [e] and [o] can also be progressive. Japanese, for example, shows glide formation following both high and mid vowels (certain conditions apply – see Kawahara 2003 for details).

(54) *Progressive glide formation in Japanese*

hea	→	he[j] a	'hair'
eakon	→	e[j]akon	'air conditioner'
siawase	→	si[j]awase	'happiness'
doa	→	do[w]a	'door'
koara	→	ko[w]ara	'koala'

guai → gu[w]ai 'condition'

Similar patterns of progressive glide formation from mid vowels can be found in Dakota (Shaw 1980) and Madurese (Stevens 1968). Again, in all of these cases, glide formation must be attributed to double association of an element to two root nodes, so a root node must be present in the empty onset in order for the process to take place.

In brief, two patterns for glide formation can be found. On the one hand, glide formation may be limited to contexts where the source vowel is simplex [i] and [u]; on the other, complex vowels such as [e] and [o] may also participate. Under the former pattern, glides are formed via double association of the root node for the source vowel. There is thus no need for the target empty onset to have its own root node. In fact, restriction of glide formation to instances where the source vowel is simplex leads inevitably to the conclusion that the target must lack a root node. Under the latter pattern, on the other hand, whenever the source is a complex vowel, glides must be formed via spreading of the select elements I and U. In this case, the target empty nucleus necessarily contains its own root node. In sum, the nature of the two patterns for glide formation is consistent with two types of target empty onset: specifically, an empty onset either with or without a root node. Importantly, these patterns are not consistent with the representational distinction that has been embraced to date by phonologists working in the GP framework, namely a distinction based on the presence versus absence of a timing slot under the empty onset.

3.6 Conclusion

What I am proposing with respect to the representation of vowel-initial syllables does not mark a radical departure from the established tradition in GP. Nonetheless, the representational distinction between empty onsets with and without a root node is crucially preferable to the GP distinction between empty onsets with and without a timing slot. First, only the proposed representational distinction is suited to account for the differing behaviour of vowel-initial versus *h-aspiré* words in French. Likewise, with respect to the various forms of glottal stop insertion and glide formation in German, Malay, Shona, and Japanese, it was shown that only a distinction based on presence versus absence of a root node accounts for the observed patterns.

In some cases, the two types of degenerate onset can simply constitute a lexical option in a language. In this case, the forms may occur in the same contexts. When the distribution of empty onsets with or without a root does not overlap, however, we find that the more degenerate forms gravitate to morphologically and prosodically recessive sites, whereas the less degenerate forms crop up in prominent sites. The representational asymmetry is thus closely paralleled by a distributional asymmetry.

The next step now is to see whether the same variation in degree of degeneracy that has been demonstrated for empty onsets is required also for empty nuclei and codas.

CHAPTER IV

EMPTY NUCLEI

4.1 Introduction

The central claim I wish to make here is that potentially two types of empty nuclei occur in phonological systems: an empty nucleus with a timing slot and a root node; and an empty nucleus with just a timing slot. These degenerate nuclei, which are either lexical or derived, can be contrasted with a lexically specified filled nucleus, as shown in Figure 4.1.

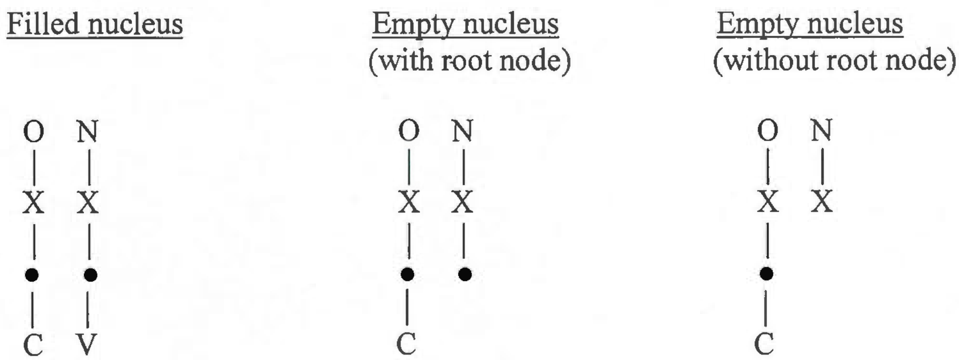


Figure 4.1 *Filled and empty nuclei*

There are two ways that empty nuclei exhibit their difference in degree of degeneracy:

i) In whether the empty nucleus may harbour a default vowel or not;
ii) In whether the empty nucleus is a potential target for local spreading of individual elements or only of entire segments.

When an empty nucleus harbours a default vowel (frequently a schwa), I consider that melodic content such as the element @ is automatically inserted into a lexically unspecified root node; in the absence of a root node, the empty nucleus fails to receive default content, and in this case, as long as the empty nucleus is not a target of spreading, it remains empty – see Figure 4.2.

<u>With a root node</u> <u>Without a root node</u>

Figure 4.2 *Schwa versus zero in empty nuclei:* C[ə] or CØ

The function of the root node in representations is to act as a docking site for subsegmental elements and as a link to the timing tier. In the absence of lexically specified elements, I contend that the root node necessarily acquires content such as default @. Of course, the fact that the empty nucleus with a root node always ends up with some form of melodic content could invite the question of whether it is not simply another form of filled nucleus. In this case, the range of representations for nuclei would be limited to two: lexically specified filled nuclei (containing schwa or any other vowel) and empty nuclei without a root node. The issue will be returned to periodically in this chapter in order to demonstrate that this apparent solution simply will not work. We have no choice but to posit more than one type of empty nucleus.

In many phonological systems that employ empty nuclei, the options in Figure 4.2 are the only ones available: either the empty nucleus receives default content or else it remains empty. In the former case, the empty nucleus contains a root node; in the latter, it lacks one. The other possibility, however, is for empty nuclei to receive content via spreading from an adjacent position. The existence of two types of spreading process provides particularly compelling evidence for a distinction based precisely on the presence versus absence of a root node (i.e., rather than on some other component of representation, such as the timing slot). As will become apparent, there is in fact a close parallel between the two types of spreading process that target empty nuclei and the two types of glide formation involving empty onsets.

One type of spreading process associated with empty nuclei goes under the name of vowel harmony. Harmony involves the spreading of individual elements to one or more successive nuclei from a neighbouring nucleus. When only a single element spreads from a complex source, as shown in, necessarily the target empty nucleus contains its own root node, or else the element will have nowhere to dock. Relatedly, it is no coincidence that filled nuclei can also participate in harmony, since these likewise provide a target endowed with a root node. In either case, the resultant configuration involves a harmonic element doubly linked to two root nodes, as shown in Figure 4.3.

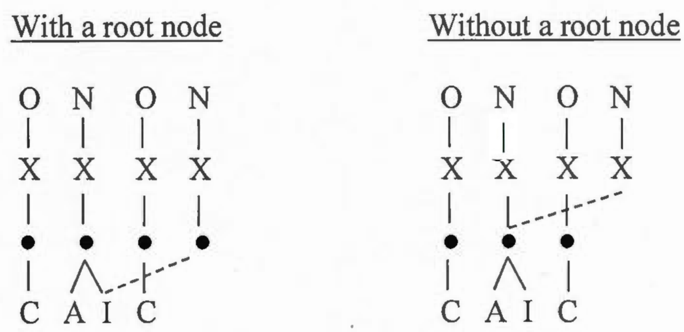


Figure 4.3 *Two types of spreading with empty nuclei as targets*

The other type of spreading process that can target empty nuclei is typically referred to as copy epenthesis. In this case, spreading is restricted to transfer of the content of an adjacent position in its entirety. To account for this restriction, the process is best depicted as involving double association of the root node to two timing slots, also shown in in Figure 4.3. Interestingly, in contrast with harmonic spreading, copy epenthesis is limited to adjacent nuclei: successive nuclei cannot contain an epenthetic copy. For reasons that remain elusive, then, it seems that single elements can spread to more than one target, whereas root nodes can be linked to timing slots under at most two nuclei.

In systems where empty nuclei can be filled via spreading, then, it is possible for empty nuclei both with and without a root node to receive phonetic expression. However, the type of content supplied via spreading depends on the degree of degeneracy of the empty nucleus. Under harmony, where the empty nucleus contains a root node, its content can be determined by spreading of single elements. In this case, the elements themselves are doubly linked to two root nodes. Consequently, the source and target of harmony may exhibit only partial identity. For example, the output of the harmonic process illustrated above in Figure 4.3 is [CeCi]. Otherwise, when the empty nucleus lacks a root node, its content can be determined only by spreading of the actual root node of an adjacent segment. In this case, the root node itself is doubly linked to two timing slots. Consequently, the source and target are always identical. Hence, the output of the spreading process shown above is [CeCe].

In GP, according to the Empty Category Principle (Kaye, Lowenstamm & Vergnaud 1990), whether an empty nucleus may remain empty or not depends upon its being licensed. The way that an empty nucleus is licensed depends upon its location (55).

(55) *Licensing of empty nuclei*

i) By parameter (domain-finally)

ii) By proper government (domain-internally or 'elsewhere')

iii) By interonset government (between an obstruent + sonorant)

iv) By 'magic' licensing (before sC-clusters)

As we will see, not all languages conform to the predicted patterns. That is, in some cases, empty nuclei in the same context (but in different lexical items) behave differently, some receiving expression, others remaining unexpressed. At times, empty nuclei are realized despite appearing in contexts where they should be licensed; and, more troubling, sometimes empty nuclei fail to be realized despite not being licensed. These facts are hard to reconcile with the Empty Category Principle. In the system I am proposing, however, whether an empty nucleus may remain empty or not depends upon its degree of degeneracy (i.e., whether it contains an unspecified root node or not). In addition, the way in which an empty nucleus may be filled is a function of its representation.

Up to a certain point, in the same way that French words may arbitrarily be consonant-initial, *h-aspiré* or vowel-initial, the selection of a filled or empty nucleus with or without a root node may at times amount to a lexical choice. In this case, the distribution of filled and empty nuclei with or without a root node is unpredictable. On the other hand, as we will see, a recurrent pattern emerges of a correlation between greater degeneracy and prosodic recessiveness. Often, only filled nuclei or, at best, empty nuclei endowed with a root node may occupy prosodically prominent sites, whereas empty nuclei without a root node are relegated to non-dominant sites.

To reiterate, GP identifies three general locations for empty nuclei: in *initial* position before sC-clusters, in *internal* position within sequences of consonants that are neither well-formed branching onsets nor coda-onsets, and in *final* position after any consonant. In the following sections, I will present evidence in favour of two types of empty nuclei, mainly in internal and final position. The evidence will come first from languages that instantiate default vowels (section 4.2) and then from languages with processes of spreading (section 4.3). The more marginal evidence I have uncovered for two types of empty nuclei in initial position before sC-clusters is presented subsequently (section 4.4). In addition, following Piggott (2003b), I present arguments for analyzing initial full and partial geminates as also preceded by empty nuclei. Throughout, the implications for the notion of the licensing of empty nuclei will be discussed. The question of whether nuclei with a default vowel such as schwa could simply be treated as another form of lexically specified nucleus will also be addressed.

4.2 Schwa-zero in empty nuclei

In this section, we consider empty nuclei that are either realized with the default vowel schwa or that remain empty. We examine evidence involving empty nuclei in medial and final position in French, German and Palauan. The occurrence of empty nuclei that behave differently in the same context requires that we posit more than one type of representation for empty nuclei, that is, with different degrees of degeneracy. Empty nuclei that are realized in contexts where they should be licensed to remain empty or, conversely, that remain empty despite not being licensed pose a serious problem for accounts based on the Empty Category Principle. The representational solution that I am proposing, on the other hand, is well-equipped to handle precisely this kind of behaviour.

4.2.1 Medial and final empty nuclei in French

A primary goal of Charette's (1991) study of vowel-zero alternation in French is to account for their distribution in terms of proper government. Still, Charette presents examples of empty nuclei in initial syllables that contravene proper government – that is, CØCV forms that are necessarily realized as CəCV. Her position is that these initial empty nuclei are exceptions to proper government.

Before examining these exceptions, we should review how proper government works. An empty nucleus is permitted to remain unexpressed if it is properly governed by an adjacent expressed nucleus to its right, as long as: i) no governing domain (i.e., branching onset or coda-onset) intervenes, and ii) no governing domain precedes the empty nucleus (as demonstrated in Charette 1990). Proper government thus applies regularly to the non-initial empty nuclei in the following forms.

(56) *Empty nuclei in non-initial syllables in French*

cad(e)nas	[kadnas]	'padlock'
mat(e)lot	[matlo]	'sailor'
pél(e)rin	[pelrɛ̃]	'pilgrim'
clav(e)cin	[klavsɛ̃]	'harpsichord'

The resultant consonant clusters [dn] [tl] [lr] [vs] are not potential branching onsets or coda-onsets, so we have little choice but to posit an intervening empty nucleus.

Curiously, empty nuclei in initial syllables are sometimes exceptions to proper government. Specifically, proper government fails to apply to initial syllables in Parisian French (PF), whether the word is (superficially) bisyllabic or polysyllabic.[53] In Quebec French (QF), on the other hand, proper government fails to apply to initial syllables in polysyllabic words only. Initial empty nuclei in 'bisyllabic' forms are properly governed in QF. Hence, in PF, initial empty nuclei are always realized as schwa; in QF, they may be realized as schwa or zero, depending on the syllable count.

(57) *Empty nuclei in initial syllables in French*

a. <u>'Bisyllabic' forms</u>	**PF → [ə]**	**QF → Ø**	
petit	[pəti]	[pti]	'small'
cheval	[šəval]	[šval]	'horse'
menotte	[mənɔt]	[mnɔt]	'handcuff'
demande	[dəmɑ̃d]	[dmɑ̃d]	'ask'

[53] By superficially 'bisyllabic', Charette means that a lexical item has only one filled nucleus in addition to the empty nucleus in the initial syllable. Whether the item does (e.g., [šəvalØ]) or does not (e.g., [pəti]) have this final empty nucleus is irrelevant. In itself, the invisibility of these empty nuclei is interesting: essentially, final empty nuclei without a root node do not count. Hypothetically, this could suggest that they are not incorporated into higher levels of prosodic structure such as the foot.

b. Polysyllabic forms	**PF/QF → [ə]**	
petitesse	[pətitɛs]	'smallness'
chevalier	[šəvalje]	'knight'
menotter	[mənɔte]	'to handcuff'
demander	[dəmãde]	'to ask'

To account for these facts, Charette (1991) attributes special status to initial syllables. The nuclei in these syllables are normally inaccessible to proper government. Exceptionally, initial nuclei in QF may be properly governed in bisyllabic forms, that is, as long as the proper governor is a stressed nucleus (stress being final in French).

Still, Charette candidly identifies three further exceptions to the exceptions in QF, namely *menu* 'menu', *femelle* 'female', and *debout* 'upright', where the initial empty nucleus must be realized despite being properly governed by a stressed nucleus: [məny], [fəmɛl], [dəbu]. Charette concedes that an explanation for these exceptional exceptions is elusive. While not necessarily fatal for her account, these outliers certainly constitute an unanswered question.

Further undermining Charette's position, in my view, are some other exceptions that she does not address. Certain polysyllabic words contain initial consonant sequences that for phonotactic reasons must be analyzed as having an intervening empty nucleus. The sequences are [pn], [pt], [gn], [mn], and [pz]/[ps]. In one set of polysyllabic words, these are produced with schwa as expected (58a); in another set, without (58b).

(58) *Initial* [pn], [pt], [gn], [mn], *and* [pz]/[ps] *in polysyllabic words*

a. With schwa		
penailler	[pənaje]	'rags' (obsolete)
petitesse	[pətitɛs]	'smallness'
guenilleux	[gəni:jø]	'tattered'
menotter	[mənɔte]	'to handcuff'

pesanteur	[pəzãtœr]	'weight'

b. <u>Without schwa</u>

pneumonie	[pnømɔni]	'pneumonia'
Ptolémée	[ptɔleme]	'Ptolemy'
gnomique	[gnomik]	'Gnomic'
mnémonique	[mnemɔnik]	'mnemonic'
psychique	[psišik]	'psychic'

Admittedly the examples without schwa are marginal (and incidentally of Greek origin), but I see no other analysis for the consonant sequences than as a pair of onsets flanking an empty nucleus. Note that they cannot be accounted for by onset-onset licensing, since this applies to an empty nucleus between an obstruent and a sonorant; the examples here include obstruent + obstruent and sonorant + sonorant. In addition, interestingly, these same sequences are all ill-formed in English and have been repaired via historical deletion of the initial consonant (as evidenced by the pronunciation of the glosses).

To reiterate, we find the following patterns and exceptions in French:

i) Differences in how an empty nucleus is realized in *initial* syllables [pətitɛs] versus *medial* syllables [kadØnas] in polysyllabic words;

ii) Differences between PF and QF in whether an empty nucleus in the initial syllable of a *bisyllabic* word is realized [pəti] or not [pØti];

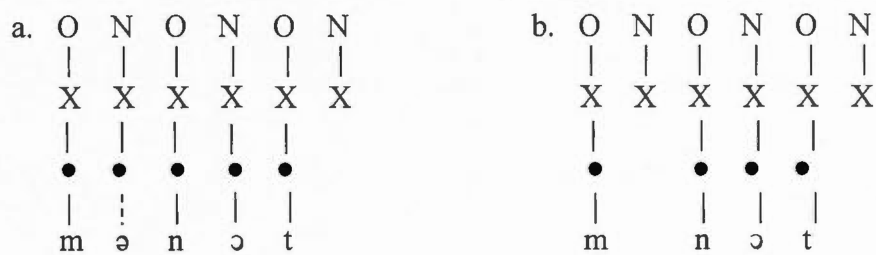
iii) Exceptions in QF to the pattern of realizing as zero an empty nucleus in the initial syllable of a bisyllabic word ([məny], [fəmɛl], [dəbu]);

iv) Exceptions to the pattern of realizing a schwa in an empty nucleus in the initial syllable of a *polysyllabic* word (e.g., [pØnømɔni]).

In short, the phonological phenomena of French are inconsistent with the notion that proper government determines whether empty nuclei are expressed or not.

The representational solution, on the other hand, does not encounter the same difficulties. Empty nuclei can be of two types: depending on whether they contain a root node, empty nuclei will or will not be realized with a schwa in French. In PF, the empty nucleus at the beginning of *menotte* [mənɔt] contains a root node that triggers schwa production (59a); in QF, the empty nucleus lacks a root node, with the result that it remains unrealized – [mØnɔt] (59b). All the same, in PF, when a vowel-final form precedes, the initial nucleus for some reason loses its root node and, as in QF, it is then realized as zero: *la menotte* [lamØnɔt].

(59) *Initial empty nucleus with (PF) or without (QF) a root node*



The same representational distinction can account for differences in how empty nuclei are realized in initial and medial syllables in polysyllabic words: [pətitɛs] versus [kadØnas]. Likewise, it can handle the pronunciation of the initial syllable of bisyllabic words like [məny] versus [mØnɔt] in QF. Similarly, the divergent realizations of the initial syllable of polysyllabic words like [pənaje] and [pØnømɔni] are amenable to a representational explanation.

Hypothetically, the capacity to represent empty categories with more than one degree of degeneracy is universal, although not all languages use this facet of the representational arsenal. Some phenomena involving empty nuclei in identical contexts in German certainly support this contention. For example, as in French, German words may begin with the consonant sequence [gn], and the intervening empty nucleus can either be realized as schwa or not: *genau* [gənau] 'precise' versus *Gnade* [gØnadə] 'grace', or *Genosse* [gənɔsə] 'comrade' versus *Gnu* [gØnu] 'gnu'.

The forms [gənɔsə] and [gØnadə], of course, also *end* in schwa. This issue will be returned to in the next section on final empty nuclei in German. First, however, we consider the realization of final schwa in Midi French.

The distinction between final empty nuclei with and without a root node is particularly apparent in Midi (Southern) French (Durand, Slater & Wise 1987; Durand 1990, 1995; Durand & Eychenne 2007; Eychenne & Pustka 2007; Eychenne 2009). Midi French has generalized schwa production at the end of words such as *porte* and *faire* that end in an orthographic 'e', but not at the end words such as *parc* and *fer* that end in a consonant. This leads to minimal pairs such as the following.

(60) *Final empty nuclei in Midi French*

a. With schwa			b. Without schwa		
tire	[tirə]	'pull'	*tir*	[tirØ]	'shooting'
faire	[fɛrə]	'to do/make'	*fer*	[fɛrØ]	'iron'
mère	[mɛrə]	'mother'	*mer*	[mɛrØ]	'sea'
roque	[rɔkə]	'castling'	*roc*	[rɔkØ]	'rock'

I analyze the words in (60a) as ending in an empty nucleus with a root node, and the words in (60b) as ending in an empty nucleus without a root node. Some varieties of Southern French are more conservative than others. These will have near categorical schwa production in empty nuclei with root nodes (whether in final or in medial position). In less conservative varieties, schwa production is more varied, but even in these a clear divide can be made between how the two types of nuclei are treated. [54]

For example, one distinction between the two types of nuclei in Midi French is that final empty nuclei without a root node trigger devoicing of a final consonant. The process applies not only to obstruents (as in Germanic and Slavic languages) but also

[54] The only context where an empty nucleus with a root node is consistently realized as zero in all Southern varieties is when the next word begins with a vowel. Hence, while *mère âgée* 'elderly mother' is pronounced [mɛraže], *mère fâchée* 'angry mother' is usually realized as [mɛrəfaše].

to sonorants. Thus, the final consonants in words such as *gaz, sud, David,* and *fer* are realized as voiceless in Midi French: [gas], [syt], [davit], [fɛχ]. The final consonants given in (60) above are thus not representative of surface forms. In Midi French, words such as *faire* and *fer*, which are homophonous [fɛʁ] in most other dialects, are distinguished not only by the production of a schwa at the end of *faire* (i.e., [fɛʁə]) but also by the devoicing of the final consonant (i.e., [fɛχ]). Interestingly, most Midi French speakers occasionally omit the final schwa in *faire* or other words, but they maintain a surface voicing distinction all the same: that is, they pronounce *faire* and *fer* as [fɛʁ] and [fɛχ]. In this case, the root node in the empty nucleus at the end of *faire* is delinked, and the reason that devoicing applies to one form and not the other is not apparent in the surface form (i.e., a case of opacity due to counterfeeding).

There is of course an alternative solution to distinguishing between nuclei that contain schwa and nuclei that remain empty in Parisian, Quebec or Midi French. This is to treat nuclei that contain schwa as not empty at all, but as in fact containing a vowel. Conceivably, words such as *tire* or *porte* that are produced with a final schwa in Midi French might end in a filled nucleus just like other vowel-final words such as *bateau* [bato] 'ship'. The same analysis might also be applied to the schwas in words such as *petitesse* or *menu* in Parisian and Quebec French. In this case, the only difference would be in the nature of the vowel appearing in the nucleus. Indeed, under this analysis, there would be no need to make the rather abstract distinction between empty nuclei with or without a root node.

One argument against such an approach is that it fails to distinguish between the full vowels of French and schwa. Schwa is not simply a vowel like any other. As emphasized in the earlier chapter on element theory, there are various aspects of schwa that make it qualitatively different from bona fide vowels and that suggest schwa constitutes more accurately the default phonetic realization of an empty nucleus (e.g., as in Anderson 1982). First, in French, schwa is the only vowel to alternate with zero. In addition, schwa differs from the full vowels of French in never bearing stress.[55] Despite stress being word-final in French (including Midi French), the final schwa in words such as *tire* or *porte* cannot carry stress.

Another argument relates to the phonetic quality of schwa in French and to alternation with zero. Although transcribed as [ə], schwa is realized phonetically as [œ] (on this issue, see, e.g., Durand 1990).[56] French also has lexical [œ]. Thus, because the schwa in the initial syllable in *genêt* 'genista' and the initial full vowel in *jeunet* 'very young' receive the same realization as [œ], the two forms are homophones when produced on their own: [žœnɛ]. However, when preceded by a vowel-final word such as *les,* the schwa deletes but not the full vowel: *les genêts* [ležnɛ] versus *les jeunets* [ležœnɛ].[57] This behaviour is more plausible if schwa is viewed as an empty nucleus with a root node that is realized as [œ], not as another full vowel which happens to have the same realization as the full vowel [œ]. Alternation with zero is a more likely property of a position with default content.

Relatedly, medial schwa in French may alternate with [ɛ] in stressed syllables, as shown in (61a), but lexical [ɛ] does not alternate with schwa in an unstressed syllable, as shown in (61b). Conversely, lexical [œ] does not alternate with [ɛ] in stressed syllables (61c).

---

[55] Exceptionally, schwa can at times be stressed in monosyllabic clitics: for instance, *fais-le* 'do it'.

[56] Although I will not pursue the issue at any length, the realization of schwa as [œ] in French may simply be a matter of phonetic implementation, That is, the elements IAU@, which are all required to specify the vowel [œ], may not actually be inserted into the surface representation.

[57] Conceivably, this contrast opens the door to an analysis of the exceptional words *menu* and *femelle* as having underlying [œ] rather than schwa in the initial nucleus (i.e., they would be like *jeunet* rather than like *genêt*). This potential solution would be more difficult to justify, however, with regard to a form such as *petitesse,* derived from *petit* (which, like *genêt,* clearly begins with a deletable schwa). It is also hard to conceive of exceptional *debout* as containing underlying [œ], since the prefix *de-* in other forms (e.g., *devant, dessous*) contains a schwa that alternates with zero. Due to the uncertainty surrounding this alternate line of analysis, then, I prefer to adhere to the established view, namely that *menu, femelle, debout,* and *petitesse* all have schwa in their initial nucleus.

(61) *Schwa-[ɛ] alternation in verb forms*

Infinitive		3rd pers sg pres		
a. Lexical schwa:				
lever	[lœve][58]	*lève*	[lɛv]	'raise'
appeler	[aple]	*appelle*	[apɛl]	'call'
peser	[pœze]	*pèse*	[pɛz]	'weigh'
b. Lexical [ɛ]:				
rêver	[rɛve]	*rêve*	[rɛv]	'dream'
mêler	[mɛle]	*mêle*	[mɛl]	'mix'
apaiser	[apɛze]	*apaise*	[apɛz]	'pacify'
c. Lexical [œ]:				
abreuver	[abrœve]	*abreuve*	[abrœv]	'water
écoeurer	[ekœre]	*écoeure*	[ekœr]	'disgust'
creuser	[krœze]	*creuse*	[krœz]	'dig'

When schwa should appear in final stressed position, the nucleus is normally realized as [ɛ] (61a). Forms such as [lɛv], I would argue, are in fact suppletive. The full vowel [ɛ] supplants schwa precisely because schwa represents a degenerate nucleus that cannot bear stress. The full vowel [ɛ], on the other hand, does not become schwa in unstressed position (61b). Likewise, the full vowel [œ] is not realized as [ɛ] in unstressed position (61c).

The implication of the exceptional behaviour of schwa outlined above is that schwa is a vocalic segment unlike any other. It is best construed not as an underlying segment but as the default realization of an empty nucleus. This reasoning brings us back full circle to the need to distinguish between those empty nuclei that are realized as schwa and those that are not. Among the former, we find final nuclei in words such as *faire*

[58] Exceptionally, I transcribe schwa as [œ] here.

in Midi French; among the latter, we find final nuclei in words such as *fer*. I propose to represent these by an empty nucleus with and without a root node respectively.

The justification for presence of a root node being associated with default schwa and its absence with zero (rather than vice versa) stems from the role the root node plays in representations. It serves as a central hub that groups together subsegmental material, relaying this material to one or more slots on the timing tier. When a position lacks lexically specified melody but is realized with a default segment such as schwa, then, it makes sense to attribute this phenomenon to the presence of a root node. Conversely, when an empty nucleus fails to trigger default melody, the absence of a root node provides a plausible explanation.

In many ways, the presence of an empty nucleus with or without a root node in a particular lexical item is simply another facet of phonological representation that can be harnessed to distinguish between lexical items. For example, in Parisian French, one dimension of difference between *penaud* [pəno] 'abashed' and *pneu* [pnø] is the presence versus absence of a root node in the empty nucleus appearing between [p] and [n]. Nothing in the context either requires or excludes the presence of a root node in the empty nucleus. Indeed, nothing requires the presence of an empty as opposed to a filled nucleus either (see *panneau* [pano] 'panel' with a filled nucleus). Similarly, in Midi French, words may end in empty nuclei either with or without a root node (*tire* vs. *tir*). In some cases, then, the choice between a filled nucleus and an empty nucleus with or without a root node amounts to a lexical option.

Nonetheless, the distribution of empty nuclei is not entirely unconstrained. We saw in (61) above that full vowels (i.e., filled nuclei) replace schwa (i.e., an empty nucleus with a root node) when the nucleus in question carries stress. It seems that empty nuclei cannot be parsed into the head of a foot in French. Apparently, absence of lexically specified elements (i.e., zero complexity) excludes a nucleus from headhood. Instead, degenerate nuclei are relegated to weak positions in the prosodic hierarchy.

Charette (1990) indicates that word-internal empty nuclei must be realized (i.e., contain a root node) when preceded by either a branching onset (*librement* [librəmã] 'freely') or a coda-onset sequence (*fortement* [fɔrtəmã] 'strongly'). Apparently, an empty nucleus without a root node is too weak a licensor for an onset that itself has licensing responsibilities. However, for reasons that remain elusive, the requirement that the empty nucleus contain a root node after a branching onset or coda-onset sequence is relaxed word-finally: the final empty nucleus remains unrealized in *libre* [librØ] 'free' and *forte* [fɔrtØ] 'strong' (fem).[59] Presumably, the pattern is related in some way to the cross-linguistic phenomenon of apocope (Piggott 1991a).

Other distributional restrictions include the fact that words in French cannot begin with a schwa (and the same applies to German but not English). Hypothetically, assuming that an initial schwa is preceded by an empty onset, there is a constraint on adjacency of degenerate constituents.

Likewise, languages generally avoid consecutive strings of empty nuclei that remain unexpressed (i.e., that lack a root node). Conceivably, this follows from a requirement that the head of a foot appear at regular intervals, so necessarily a filled nucleus appears to serve this function. A case in point is Moroccan Arabic (62) (Mohamed Guerssel, personal communication).

(62) *Alternation between realized and unrealized nuclei in Moroccan Arabic*

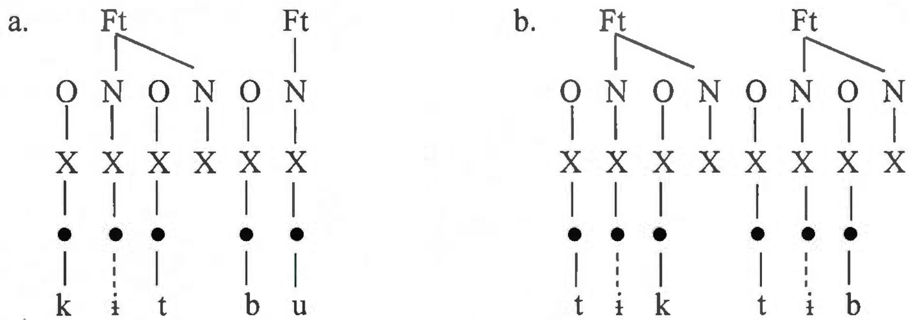
kØtØbØ	→	[ktɨb]	'he wrote'
kØtØb-u	→	[kɨtbu]	'they wrote'
tØ-kØtØbØ	→	[tɨktɨb]	'you (masc.) write'

[59] Interestingly, unexpressed final empty nuclei can license a branching onset in French (e.g., [tablØ] 'table') but not in English (e.g., [tejbəl]/*[tejblØ]). Nikiema (1989, cited in Charette 1991) attributes this variation to a direct/indirect licensing parameter. In languages like French, where the parameter is set at 'on', final (but not medial) empty nuclei that remain empty are 'indirect licensors'. That is, they can license an onset head to govern a dependent even when not strictly adjacent.

tØ-kØtØb-i → [tkɨtbi] 'you (fem.) write'

The underlying forms above contain as many as four empty nuclei in a row (CØCØCØCØ), but the surface forms always alternate between realized and unrealized nuclei. In GP, this kind of pattern is accounted for via the notion of proper government. Equally, it can be accounted for by positing a binary trochaic foot, with a constraint that the head of the foot must contain either a filled nucleus or an empty nucleus with a root node (or, conversely, the dependent nucleus in the foot must lack a root node). These patterns are illustrated below for [kɨtbu] and [tɨktɨb].[60]

(63) *Foot structure in Moroccan Arabic*



Given a trochaic foot constructed iteratively from the right edge, with filled nuclei automatically constituting the head of a foot and feet being preferably binary, these are exactly the patterns we would expect to find and not, say, *[ktɨbu] instead of [kɨtbu] or *[tɨktɨbi] instead of [tkɨtbi]. Note that the correct forms can be generated whether we assume that all of the empty nuclei underlyingly contain or lack a root node. In the former case, a process of lenition would remove the root node from empty nuclei in dependent sites; in the latter, a process of fortition would add a root node to empty nuclei in head position. Importantly, one consequence of this account

[60] The foot patterns for [ktɨb] and [tkɨtbi] are a bit more complicated to capture due to the initial unrealized nucleus, which is necessarily an unstressed dependent. This extra nucleus could, however, be incorporated at the level of the 'superfoot' (Davis & Cho 2003) or simply of the prosodic word.

is that there is no need for recourse to proper government to determine the distribution of realized and unrealized empty nuclei. As we will see next, however, while an analysis of vowel-zero alternation based on foot structure is promising, it is not without its own problems.

Proper government is also used in GP to account for the various patterns of schwa-zero alternation involving empty nuclei in clitics combined with lexical forms. For example, in *la queue de ce renard* 'the tail of this fox', the clitics *de ce* and the word *renard* all contain an empty nucleus. These empty nuclei may each remain empty, as long as no two empty nuclei in a row are unrealized. Hence, [dØsərØnar] or [dəsØrənar] are possible (among other combinations), but *[dəsØrØnar] or *[dØsØrənar] are not. Certainly, since proper government of empty nuclei only applies when the governor itself is realized, this would seem to explain why the former are acceptable but not the latter, as illustrated in (64a,b).

(64) *Proper government of empty nuclei in clitics and lexical forms*

```
a. O N O N O N O N        b. O N O N O N O N
   | | | | | | | |           | | | | | | | |
   X X X X X X X X           X X X X X X X X
   | | | | | | | |           | | | | | | | |
   d Ø s ə r Ø n a r         d ə s Ø r Ø n a r
```

Clearly, proper government 'works' to account for the patterns in (64a,b). On the other hand, where it does not work is when *renard* appears on its own. As mentioned previously, in standard Parisian French, an initial empty nucleus in a bisyllabic form is not licensed to remain empty unless the preceding word ends in a vowel. Hence, *renard* must be realized as [rənar], but *ce renard* can be realized as [sərØnar]. A preceding realized nucleus is thus required in order to permit proper government by a following nucleus.

The role played by the preceding nucleus suggests that we are once again dealing with a phenomenon determined by trochaic foot structure. Nonetheless, this explanation is problematic, since it would require that foot structure be determined after clitics are attached. In addition, it would need to apply also after phrases are formed, since the empty nucleus in *renard* can remain empty not only when a clitic precedes, but also when a vowel-final lexical word such as *beau* precedes: *beau renard* [borØnar].

In sum, neither proper government nor an analysis based on foot structure is entirely successful in capturing the cross-linguistic patterns of alternation in the realization of empty nuclei. Indeed, further cases where proper government runs into trouble will be presented in sections 4.2.3 (on Palauan) and 4.3.2 (on Turkish). I conclude then that further research on this issue is required, apparently necessitating an innovative solution that is yet to be imagined.

4.2.2 Final empty nuclei in German

German has words such as *Tag* [tak] 'day' and *Bad* [bat] 'bath' that are consonant-final. In my view, these final consonants are onsets followed by an empty nucleus without a root node. Due to the absence of a root node, the nucleus fails to be realized. Interestingly, as in Midi French, a final empty nucleus without a root node triggers onset devoicing in German. German also has words that end in schwa, however, including the examples encountered earlier such as *Gnade* [gØnadə] and *Genosse* [gənɔsə]. These end in an empty nucleus with a root node that triggers default content. This section examines in depth just such examples from German of final empty nuclei with different degrees of degeneracy.

Certain forms in German (Giegerich 1985; Wiese 1986, 2000; Hall 1989; Brockhaus 1995) pose a problem for the GP view that final empty nuclei are licensed by parameter. Under the GP view, final empty nuclei should either remain empty in a language or they should be realized by a vowel determined by default (or else by local spreading), depending on the parameter setting. The problem with German is that, like Midi French, it has both consonant-final and schwa-final forms. Consonant-final forms indicate that the parameter licensing final empty nuclei is set at 'on'. On the other hand, as long as schwa constitutes the default realization of an empty nucleus in German, schwa-final forms suggest the parameter for final empty nuclei should be set at 'off'.

A schwa that appears in final position in German may be part of the root, although more commonly it constitutes a separate suffix. Some examples of noun roots with final schwa are provided below.

(65) *Final schwa in German noun roots*

Singular		
Ware	[va:rə]	'commodity'
Hose	[ho:zə]	'pants'
Gasse	[ga:sə]	'alley'
Mieke	[mikə]	person's name
Rolle	[rɔlə]	'role'
Warze	[vartsə]	'wart'

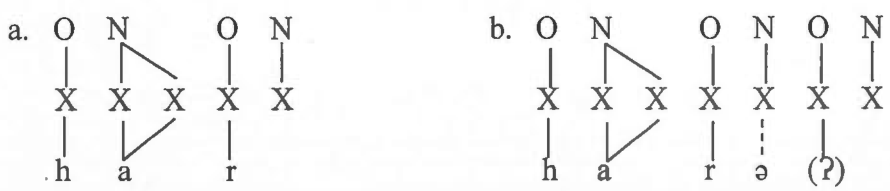
A final schwa can also be used as a multifunction inflectional suffix that attaches to nouns, adjectives and verbs alike. (Indeed, intriguingly, schwa is the *only* vowel that appears in inflectional affixes in German.) For example, the plural of the following consonant-final words is in each case formed by realizing a final schwa.

(66) *Suffixal schwa in German nouns*

Singular		Plural		
Haar	[ha:r]	*Haare*	[ha:rə]	'hair'
Gas	[ga:s]	*Gase*	[ga:zə]	'gas'
Mass	[ma:s]	*Masse*	[ma:sə]	'measure'
Zweck	[tsvɛk]	*Zwecke*	[tsvɛkə]	'purpose'
Bauch	[baux]	*Bäuche*	[bɔjçə]	'belly'
Kreuz	[krɔjts]	*Kreuze*	[krɔjtsə]	'cross'

Working within the GP framework, Brockhaus (1995) addresses the issue of German words ending in both consonants and schwa. In her view, schwa is the realization of an unlicensed empty nucleus. As is customary in GP, she assumes domain-final nuclei to be licensed to remain empty by parameter. Word-internal empty nuclei are licensed by proper government. The challenge then is to account for the occurrence of both *Haar* [ha:r] and *Haare* [ha:rə]. Brockhaus resorts to a complex representation, based on a proposal by Kaye (1990) for a similar phenomenon in Turkish. This involves an empty nucleus followed by another empty nucleus (along with an intervening glottal stop that never surfaces) (67b).[61]

(67) *Brockhaus's representation of final schwa:* Haar *versus* Haare



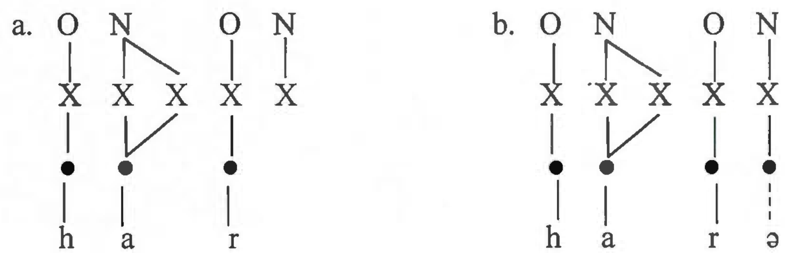
The empty nucleus following [r] in (67a) is licensed to remain empty by virtue of being final: hence, [ha:r]. In (67b), on the other hand, this same nucleus is no longer

[61] This is in fact only the ultimate stage in Brockhaus's derivation of forms such as *Haare.* The full account starts with a total of three empty nuclei at the end of the word, so the full picture is more complicated than what is shown here. The reader is referred to the original for details.

final, nor is it properly governed. Hence the nucleus must receive phonetic realization: [ha:rə]. Clearly, Brockhaus's account generates the surface output, but the complexity of the account undermines its appeal and poses a problem for learnability.

The solution advocated here is much simpler: German words may end in empty nuclei with or without a root node (68).

(68) *Proposed representation of final schwa:* Haar *versus* Haare



Under this view, when the plural is formed with a suffix schwa, the process involves addition of a root node to the final empty nucleus, thus triggering the production of schwa. The plural morpheme is thus made up only of an unspecified root node.[62] This strikes me as a more plausible (and learnable) process than what Brockhaus posits.

An alternative solution might be to represent schwa in German as a full vowel with lexically specified content. This alternative was shown to be untenable for French, and I would argue that it cannot apply to German either. First, whether final or medial, schwa in German cannot carry stress. Other final vowels in German, on the other hand, have no problem being stressed: for example, *Radau* [ra'dau] 'noise' and *Schweinerei* [švainə'rai] 'filth'. Schwa is also the only vowel that alternates with zero. For example, the phrase *ich habe* 'I have' can either be realized as [ıçØhab**ə**] or as [ıçØhap**Ø**]. In sum, schwa behaves like a degenerate vowel, not like a full vowel, so it seems justified to maintain a distinction.

---

[62] There are other means of forming the plural of nouns in German: alongside [ə], the suffixes [ən], [ər], and [(ə)s] (with or without umlaut) may be used (Fox 2005).

So far we have seen that a representational account based on different degrees of degeneracy in empty nuclei is more effective in handling the distribution of schwa and zero in medial and final position in Parisian, Quebec and Midi French, as well as in German. Further evidence for the proposed analysis comes from phenomena involving unstressed vowels in the Austronesian language Palauan.

4.2.3 Medial and final empty nuclei in Palauan

Depending on the context, unstressed full vowels in Palauan alternate with either zero or schwa (Flora 1974; Josephs 1975, 1990; Zuraw 2003). In final position, unstressed vowels undergo deletion (69) (accents indicate stress).

(69) *Deletion of final unstressed vowels*

badu	→	báθ	'rock'
kɛri	→	kɛ́r	'question'
mada	→	máθ	'eye'

In internal position, on the other hand, unstressed vowels generally reduce to schwa. When the same roots as in (69) appear with a suffix, the vowels [u], [i] and [a] that are deleted above attract stress and surface as is – see (70). However, the unstressed vowels [a] and [ɛ] in the initial syllable are now reduced to schwa.

(70) *Reduction to schwa of internal unstressed vowels*

badu-k	→	bəðúk	'my rock'
kɛri-k	→	kərík	'my question'
mada-k	→	məðák	'my eye'

Next, when the same roots appear with stressed suffixes, both internal vowels are reduced to schwa. Some examples with the root [badu] are given below in (71).

(71) *Reduction to schwa before a stressed suffix*

badu-mam	→	bəðəmám	'our (excl.) rock'
badu-miw	→	bəðəmíw	'your (pl.) rock'
badu-ɾiɾ	→	bəðəɾíɾ	'their rock'

From a GP perspective, it appears that full vowels in Palauan lose their segmental content and become empty nuclei when unstressed.[63] The fact that word-final empty nuclei remain empty can be attributed to the final empty nucleus parameter being set at 'on' (72a). Proper government, on the other hand, appears to be inactive in Palauan, since word-internal empty nuclei must be realized with default schwa, even when they are potentially properly governed by a neighbouring nucleus (72b).

(72) a. *Final empty nucleus*
(parametrically licensed)

```
O N O N
| | | |
X X X X
| | |
b á θ
```

b. *Internal empty nucleus*
(unlicensed by proper government)

```
O N O N O N
| | | | | |
X X X X X X
| : | | |
b ə ð ú k
```

The situation is more complicated than it appears, however. Word-final empty nuclei are sometimes realized by schwa, and word-internal empty nuclei are sometimes not realized. Specifically, word-final empty nuclei are realized by schwa following two consonants (73). In GP terms, the final empty nucleus apparently is not licensed to remain empty under these conditions, despite the appropriate parameter setting. Hypothetically, the final empty nucleus might need to be realized in order to properly

[63] Note that the element content for schwa (i.e., @) is not necessarily present in the full vowel counterparts, so reduction really does involve total element suppression and default insertion of @, not simply selective partial removal of elements.

govern a penultimate empty nucleus between the two final consonants (e.g., [pØsípØsə]).

(73) *Final schwa*

bsibs	→	psípsə	'drill'
kbokb	→	kpókpə	'wall'
ralm	→	rálmə	'water'
dɛŋʔokl	→	dɛŋʔóklə	'sit'

Conversely, the forms in (74) start with consonant sequences that, according to the complexity condition, are impossible branching onsets.

(74) *Internal empty nuclei in initial consonant sequences*

tmák	'together'
pŋák	'my flower'
ptúʔ	'star'
tpák	'my spit'
kpókpə	'wall'

Even disregarding the question of elemental complexity, the presence of the sequences [pt] and reverse-order [tp] sends a clear signal that these are not branching onsets – phonotactic constraints usually impose a strict order on branching constituents. Instead, then, the consonants are best analyzed as a pair of onsets separated by an empty nucleus that remains empty. This is problematic because the internal empty nucleus appears to be licensed by proper government to remain empty here, although we have already seen that proper government does not apply elsewhere in Palauan.[64] Furthermore, because the consonants are not limited to obstruent-sonorant pairs, interonset licensing cannot be invoked.

To further complicate matters, word-internal unstressed [u] may be deleted rather than realized as schwa (75).

(75) *Deletion of word-internal* [u]

ul-bunut	→	ulbúnt	'drill'
dakul-iy	→	doklíy	'sit'

Note that the form [ul-bunut] → [ulbúnØtØ] ends in two unexpressed empty nuclei: inexplicably, the penultimate empty nucleus remains empty even despite its not being properly governed.

In sum, there is a general tendency in Palauan for final empty nuclei to remain unrealized and for internal empty nuclei to be realized as schwa. Nonetheless, this pattern is by no means constant: in fact both final and internal empty nuclei may be realized as either schwa or zero. This situation poses a serious problem for the GP position that empty nuclei must be licensed to remain empty. Palauan has empty nuclei that must be expressed despite being parametrically licensed (e.g. [psípsə̲]) or potentially properly governed (e.g. [bə̲ðúk]). It also has empty nuclei that remain empty despite being unlicensed (e.g. [ulbún<u>Ø</u>tØ]). Using orthodox GP tools, then, the patterns of realized and unrealized empty nuclei in Palauan simply cannot be captured.

The representational approach I am advocating, on the other hand, provides a solution to the Palauan paradox. Two types of empty nuclei can appear in Palauan forms: i)

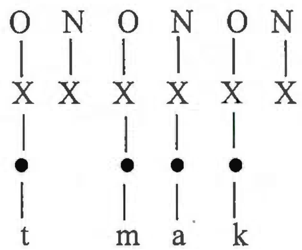
---

[64] Admittedly, I am not comparing identical cases: the initial empty nucleus is in a form such as [bəðúk] is derived through a process of vowel reduction, whereas the initial empty nucleus in [tØmák] is underlying. Still, I cannot see why this should make any difference for whether the nucleus is realized or not. Certainly, the standard position in GP is that processes apply whenever the conditions for their application are met. The ultimate conditions in [bəðúk] and [tØmák] are the same: an empty nucleus appears in a context where it is potentially properly governed.

empty nuclei with both a timing slot and a root node (though no segmental content); and ii) empty nuclei with just a timing slot. The former are realized by a default schwa; the latter fail to be realized. Both may appear in word-final or internal position, and they may be either lexically specified or generated at the surface.

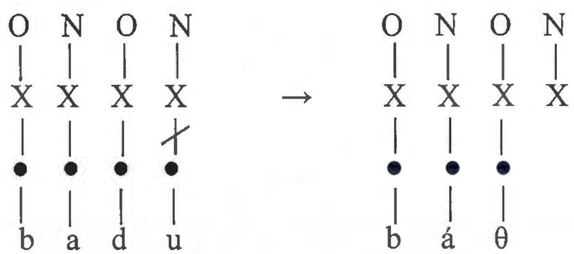
For example, both the internal and final empty nuclei in the form [tmák] contain only a timing slot (24). Consequently, they serve the purpose of licensing the preceding onset, but fail to license segmental content themselves. Both are lexically specified (i.e., not derived via reduction).

(76)

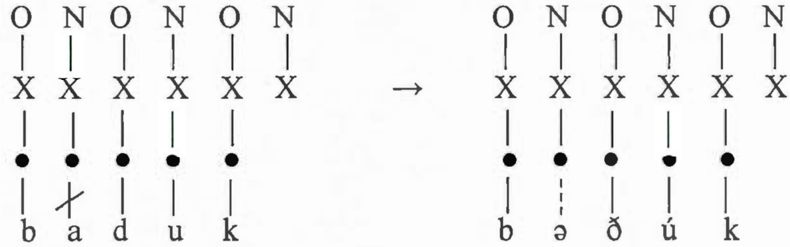


The final empty nucleus in [báθ] also contains only a timing slot and hence fails to be realized. In this case, however, the empty nucleus is generated by delinking the root node and associated segmental material from lexical [badu] (77).

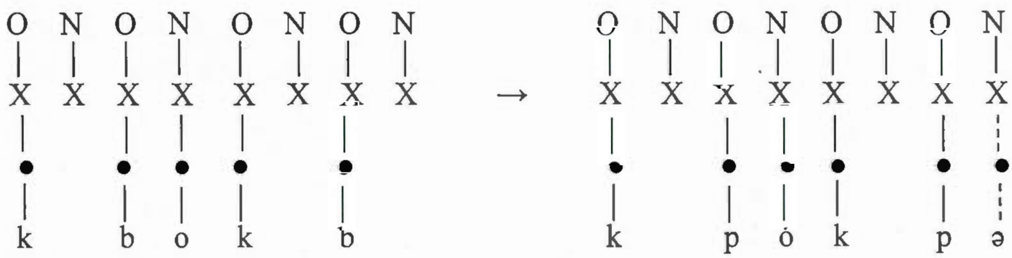
(77)



The same process of root-node delinking applies to word-internal [u] in [ul-bunut] → [ulbúnØt]. Consequently, [u] is deleted. More commonly, however, word-internal unstressed vowels reduce to schwa. In such cases, only the segmental content associated with the root node is delinked. This is the process that applies in [baduk] → [bəðúk] (78).

(78) 

The empty nuclei with a root node in [bəðúk] and without a root node in [báθ] are the result of a lenition process applying to an unstressed full vowel. The empty nucleus with a root node at the end of [kpókpə], on the other hand, is the result of a fortition process inserting a root node into the empty nucleus after two final consonants (79).[65]

(79) 

In sum, I am arguing that Palauan employs three representations of nuclei. A filled nucleus contains a full vowel that is realized if it appears in a surface form. An empty nucleus without a root node fails to be realized in a surface form. An empty nucleus with a root node, on the other hand, is realized with default content. In Palauan, the empty root node automatically receives the element @, resulting in schwa production. In GP, a nucleus with schwa is of course simply an empty nucleus that fails to be licensed to remain empty. This analysis has been shown to be inadequate to account for the distribution of schwa and zero in Palauan.

[65] The proposed fortition and lenition processes may strike the reader as a bit of a sleight of hand, but processes of addition or subtraction of elements of prosodic structure are well-established. For example, instances of vowel-lengthening or shortening are usually captured by insertion or deletion of a timing slot. Selayarese (Mithun & Basri 1986) provides a notable example. In this language, stressed vowels must be long, and usually they occur in penultimate position such that, when a suffix is added, the stressed vowel (and vowel length) alternates: see [gó:lo] 'ball' but [goló:ku] 'my ball'. By extension, I assume that languages may employ processes of insertion or deletion of a root node.

We have already considered an alternative to the solution I am advocating involving degrees of degeneracy of empty nuclei: schwa might be represented as simply a filled nucleus. Under this analysis, we could reduce our representational set to two: filled nuclei and empty nuclei without a root node. This view fails to account for the special properties of schwa, however. In Palauan, as in the previous languages we considered (French, German), schwa only appears in unstressed nuclei. In addition, schwa can occur as the reduced form of any vowel in the Palauan inventory [i, ɛ, u, o, a]. As pointed out in note 63, however, the element @ is not actually present in most of these vowels (only in [ɛ]), so schwa cannot be posited as the output of a reduction process per se. Instead, I propose that Palauan has two types of deletion process. One removes only segmental content but leaves the root node in place. This is what happens when an unstressed vowel reduces to schwa. The other process delinks the root node and associated content of an unstressed vowel. This results in zero. In addition, empty nuclei in Palauan may be lexically specified either with or without a root node.

To push the analysis further, it is desirable to consider whether the occurrence of the two types of empty nuclei, representing different degrees of degeneracy, is not in some way constrained. Clearly, when empty nuclei both with and without a root node appear in common contexts, no restrictions are in place, and this is certainly by and large the case in Palauan. Nonetheless, one pattern emerges involving the reduction of full vowels: in final position, these reduce to zero; in internal position, they reduce to schwa. The latter fact is of little interest, since word-internally both schwa and zero are found. That is, while the observed process is [baduk] → [bə̲ðúk], nothing appears to exclude [baduk] → [bØ̲ðúk] as potential process. In final position, however, zero is clearly preferred, schwa occurring only under special conditions (when preceded by a series of two consonants). This process of apocope is common cross-linguistically (for further examples, see, e.g., Piggott 1991a). I am not sure why there should be a preference for this more degenerate nucleus at the right edge of words, beyond that, by ending a lexical item phonetically with a consonant, particularly when this is preceded by a stressed syllable, an effect of edge demarcation is achieved, which might be an aid to parsing (Kaye 1989).

In French, German and Palauan, the element @ appears automatically in an empty nucleus with a root node, whereas empty nuclei without a root node remain unexpressed. The other possibility is for an empty nuclei to be filled by local spreading. Interestingly, there are two ways for spreading to operate, depending on whether the target is an empty nucleus with a root node or without a root node. The next section examines both forms of spreading in empty nuclei, in Selayarese and Winnebago on the one hand, and Turkish on the other. Turkish proves to be a particularly interesting case because we find both default and harmonic vowels in empty nuclei. In addition, we find empty nuclei that fail to be realized. The need to distinguish between two types of empty nuclei is thus strongly motivated in Turkish.

4.3 Local spreading in empty nuclei

Local spreading can target either nuclei with full vowels or empty nuclei. When full vowels are the target, an element is simply added to the elements that are already there. For example, earlier we saw examples of laxing harmony in Tangale and Akan – see (26) and (27) in Chapter II. In these cases, the element @ spreads from one nucleus to subsequent nuclei, triggering laxing of the vowel that is already present. When the content of the nucleus is determined entirely by the source, on the other hand, the target is an empty nucleus. In this case, there are two possibilities. Either the source always shares all of its content with the target or else only select elements from the source are shared. Under the first pattern, I analyze the target as an empty nucleus without a root node: the source vowel spreads in its entirety via double association of its root node to the timing slot in the target. As a result, source and target are necessarily identical. Typically, this phenomenon is referred to as copy epenthesis. Under the second pattern, I argue that the target empty nucleus must contain its own root node. Consequently, it can harbour select individual elements from the source, similar to when the target is a full vowel. As a result, source and target are not necessarily identical. Typically, this phenomenon is referred to as vowel harmony.

The following section first revisits some data from Selayarese that were introduced previously in Chapter I. Selayarese shows local spreading in empty nuclei without a root node (copy epenthesis). Winnebago provides further examples of this type of spreading. Subsequently, we examine extensive data from Turkish. This language is of particular interest because it instantiates empty nuclei both with and without a root node. In Turkish, local spreading targets only empty nuclei with a root node (vowel harmony); when the empty nucleus lacks a root node, it remains empty. Harmony also targets full vowels, but only those lexically specified for the element A.

4.3.1 Local spreading in empty nuclei without a root node

Selayarese words that end in one of the consonants [r, l, s] cannot surface as is; instead they appear with a final epenthetic vowel that is identical to the vowel in the preceding nucleus (80) (Mithun & Basri 1986).

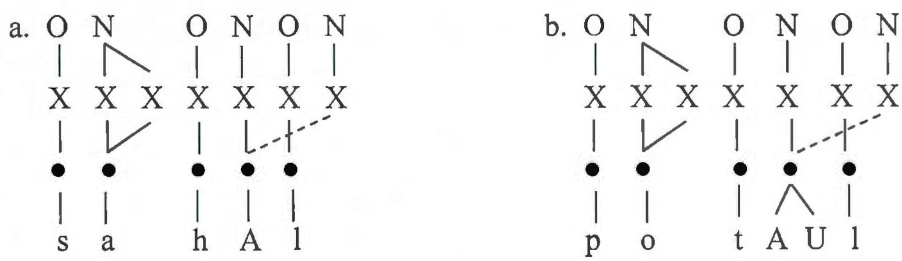
(80) *Vowel epenthesis in Selayarese*

sahal	→	[sá:hala]	'profit'
lamber	→	[lámbere]	'long'
potol	→	[pó:tolo]	'pencil'
sussul	→	[sússulu]	'burn'
kikir	→	[kí:kiri]	'metal file'

The stress patterns that are instantiated in these words indicate that the final vowel is not lexically specified: normally, stress is penultimate in Selayarese, but here stress falls on the antepenultimate nucleus. Apparently, when it comes to stress assignment, the final vowels in these forms do not count. That stress truly is on the antepenultimate syllable in these forms is confirmed by the process of vowel lengthening which applies if the syllable is open.

Under my analysis, the forms in (80) all end in an empty nucleus without a root node. The final vowel is determined by spreading from the root node associated to the preceding nucleus, creating a doubly linked structure. As a result, the content of the two nuclei is identical. The process is illustrated below for [sa:hala] and [po:tolo].

(81) *Local spreading to empty nuclei without a root node*



Another way of generating the final vowels in the forms in (80) might be to treat them as so-called copies of the preceding vowels (e.g., Kitto & de Lacy 1999). Under this view, rather than spreading, the element A in (81a) and the elements AU in (81b) would be repeated in the final nucleus, although they would not be lexically specified. In other words, the final nuclei would start off as empty, but then be supplied with their own content based on the content of the preceding nucleus. In the surface form, then, the final nuclei would have a representation identical to the previous nucleus.[66]

This copy-vowel solution is less desirable than the local spreading account for a number of reasons. First, if the surface representation of the final nucleus is identical

[66] One advantage to this approach is that it avoids the apparent violation in (81) of the No Crossing Constraint. Nonetheless, Kawahara (2004) presents a number of arguments in favour of viewing copy (or 'echo') epenthesis as in fact a case of spreading. In addition, various solutions have been proposed to account for such apparent crossing violations (e.g., Clements & Hume 1995).

to the preceding nucleus (i.e., both are filled nuclei dominating their own elements), it is arguably more difficult to see why this final nucleus is overlooked in stress assignment. Next, in Selayarese, final vowels in empty nuclei alternate with zero before vowel-initial suffixes, but full vowels do not. Hence, [lámbere] 'long' (with a final empty nucleus) is realized as [lámber-aŋ] 'longer' before the suffix [-aŋ], but [lóhe] 'many' (with a lexically specified final vowel) is realized as [lóhe-aŋ] 'more'. If the final vowel in [lámbere] is endowed with its own content, however, there is nothing to distinguish it from the final vowel in [lóhe]. There is no clear motivation for this vowel to be deleted before a vowel-initial suffix. If the final vowel in [lámbere] is determined by spreading, on the other hand, its representation diverges from the final vowel in [lóhe], and differences in behaviour are to be anticipated.

Another way in which Selayarese final vowels in empty nuclei diverge from final specified vowels is in how they behave before certain consonant-initial suffixes. For example, when the suffix [-mu] 'your' (fam) is added to a form, stress shifts to the preceding nucleus. When this nucleus contains a lexical vowel, as in [sahala] 'sea cucumber', the vowel lengthens: [sahal**aː**mu]. When the nucleus is underlyingly empty, on the other hand, as in [sahal] 'profit', the vowel does not lengthen. Instead, a glottal stop is added in coda position: [sahal**aʔ**mu]. If this final vowel is a copy, it makes little sense that it should be unable to lengthen. On the other hand, if it is determined by spreading from the preceding nucleus, lengthening would necessitate an additional association line from the source root node. It is precisely this complex process that appears to be illicit in Selayarese. This restriction is not surprising, given that one particularity of processes that target empty nuclei through spreading of the root node is that they are limited to immediately adjacent targets. They cannot target successive empty nuclei, such that multiple association of the root node is apparently never tolerated.

Winnebago provides another example of empty nuclei that lack a root node and that are filled by local spreading (Miner 1979, 1989; Hale & White Eagle 1980; Alderete 1995; Davis & Baertsch 2012). The context for empty nuclei to be filled in Winnebago is defined by what Miner (1979) terms Dorsey's Law: an epenthetic vowel identical to the following vowel appears between a cluster comprised of a voiceless obstruent and a sonorant (82a). Note that when a cluster other than voiceless obstruent + sonorant appears, no epenthetic (i.e., harmonic) vowel occurs: for example, when the obstruent is voiced (82b) or when an obstruent appears in second position (82c).

(82) *Vowel epenthesis in Winnebago*

a.	hipres	→	[hiperes]	'know'
	krepna	→	[kerepana]	'unit of ten'
	š-rugas	→	[šurugas]	'tear' (2nd pers)
b.	haracab-ra	→	[haraca**br**a]	'the taste'
c.	sgaa	→	[**sg**aa]	'white'
	kšee	→	[**kš**ee]	'revenge'
	račgã	→	[ra**čg**ã]	'to drink'

As with Selayarese, vowel epenthesis can be viewed as spreading of the vowel from an adjacent nucleus. Since it is the entire vowel that spreads, including complex vowels such as [e], this suggests the process involves double association of the root node dominating the vocalic elements. Under these circumstances, we can again posit a target empty nucleus that lacks a root node.

Not all languages with local spreading to fill empty nuclei function like Selayarese and Winnebago. In some cases, only select elements spread from one nucleus to another, rather than the entire vowel. Under these circumstances, the target empty nucleus must contain its own root node in order to harbour the individual elements. A noteworthy example of such a language is Turkish.

4.3.2 Harmony in empty nuclei with a root node

The phonological system of Turkish is well documented. General introductions include Lees (1961), Lewis (1967), Underhill (1976), Comrie (1997), and Göksel and Kerslake (2005). Other seminal texts focus on issues such as vowel harmony (Yavas 1980; Clements & Sezer 1982), compensatory lengthening (Sezer 1986; Kornfilt 1986), and stress (Inkelas 1999; Kabak & Vogel 2001). More recently, vowel harmony has been re-analyzed in terms of OT/Correspondence Theory (Krämer 2003) and in terms of the locality conditions that determine a possible harmony system (Nevins 2010). I am interested in Turkish vowel harmony from the perspective of empty nuclei and element theory. Ground-breaking work in this vein appears in van der Hulst and van de Weijer (1991), Charette and Göksel (1996), and Charette (2006).

Conventionally, Turkish is described as having back and round harmony. That is, all vowels in a harmonic form are either back or front and either round or unround. Under element theory, Turkish vowel harmony involves spreading of the elements I and U. When I appears in the initial nucleus of a form and spreads to subsequent nuclei, all vowels in the form are front; in the absence of I, all vowels are back. Similarly, U-spreading from the initial nucleus triggers rounding of subsequent vowels; absence of U results in unround vowels. Spreading of I and U combined triggers front round vowels.

I-harmony applies both to empty nuclei and to nuclei that are lexically specified for the element A. U-harmony applies only to empty nuclei. Turkish has two types of empty nuclei: those with a root node, which are potential targets of harmony, and those without a root node, which do not undergo harmony. Empty nuclei without a root node are transparent to harmony, allowing elements to spread across them to an adjacent nucleus. A-nuclei, on the other hand, are opaque to U-harmony: they block the progression of U to a subsequent potential target. Finally, the A-element does not participate in harmony (i.e., it does not spread). Examples will be provided forthwith, but first the Turkish vowel inventory will be presented.

Turkish is usually described as having the following eight vowels (Table 4.1).

Table 4.1 *Turkish vowel inventory*

Front				**Back**			
Unround		Round		Unround		Round	
[i]	(I)	[ü]	(I U)	[ɨ]	(@ I)	[u]	(U)
[e]	(I A)	[ö]	(I U A)	[a]	(A)	[o]	(U A)

The vowel [ɨ], however, is special. In my view (shared by van der Hulst & van de Weijer 1991, Charette & Göksel 1996, and Charette 2006), [ɨ] is the default realization of an empty nucleus (i.e., with a root node, in my system). Lexically, then, any nucleus realized as [ɨ] has no underlying element content, but the elements @I are automatically inserted into the surface form. An empty nucleus is realized as [ɨ] only when it does not undergo harmony (i.e., when the preceding nucleus lacks I and U). Empty nuclei without a root node do not participate in harmony nor do they receive default [ɨ]. On the other hand, nuclei with [a] (i.e., only the A-element) are normally targets of I-harmony (cf. Clements & Sezer 1982 for a few exceptions).

Harmony in Turkish operates from the first vowel in a form and targets subsequent vowels in the root and in any suffixes, which are numerous (though Turkish lacks prefixes). Examples of root harmony will now be presented (based on van der Hulst

& van de Weijer 1991), followed by examples of suffix harmony. First, I-harmony targets both empty nuclei (with a root node)[67] and A-nuclei, as shown in (83).

(83) *I-harmony in Turkish roots*

Target: Empty nucleus (w/ RN)		Target: A-nucleus	
i - Ø → i - i	e - Ø → e - i	i - A → i - e	e - A → e - e
kiši 'person'	degis 'change'	ince 'thin'	kere 'time'
gibi 'like'	yedi 'seven'	igne 'needle'	gebe 'pregnant'
inci 'pearl'	eski 'cold'	dičer 'other'	tepe 'hill'

Next, U-harmony targets empty nuclei but not A-nuclei (84).

(84) *U-harmony in Turkish roots*

Target: Empty nucleus (w/ RN)		Target: A-nucleus	
u - Ø → u - u	o - Ø → o - u	u - A → u - a	o - A → o - a
kuru 'dry'	soguk 'cold'	tuhaf 'strange'	boga 'bull'
ugur 'fortune'	yorgun 'tired'	bugday 'wheat'	oda 'room'
nutuk 'speech'	oku 'read'	muhak 'new moon'	dogar 'is born'

In addition, I- and U-harmony can combine when the target is an empty nucleus, but only the I-element participates when the target is an A-nucleus (85).

(85) *Combined I- and U-harmony in Turkish roots*

Target: Empty nucleus (w/ RN)		Target: A-nucleus	
ü - Ø → ü - ü	ö - Ø → ö - ü	ü - A → ü - e	ö - A → ö - e
ütü 'iron'	sögüt 'willow'	dümen 'wheel'	öyle 'thus'
üzüm 'grape'	gönül 'heart'	dügme 'button'	gönder 'send'

[67] For brevity's sake here, I will not always state that the empty nucleus contains a root node. It should be understood, however, that empty nuclei in Turkish that are targets of harmony have a root node.

čünkü 'because'	dövüs 'fight'	müspet 'proven'	köpek 'dog'

Finally, when the first nucleus is empty or contains only the A element, harmony is inoperative. Under conditions where the context for harmony is missing, empty nuclei are realized as default [ɨ], and A-nuclei as [a] (86).

(86) *Context for harmony missing in Turkish roots*

Target: Empty nucleus (w/ RN)		Target: A-nucleus	
Ø - Ø → ɨ - ɨ	A - Ø → a - ɨ	A - A → a - a	Ø - A → ɨ - a
kɨsɨm 'part'	altɨ 'six'	kara 'black	gɨrtlak 'throat'
sɨnɨr 'border'	yalɨ 'villa'	tavšan 'rabbit'	hɨtta 'province'
sɨgɨr 'sick'	kadɨ 'judge'	hasta 'sick'	kɨyak 'excellent'

The same patterns appearing in roots are found in suffixes. Suffixes may contain an empty nucleus as in *-(s)Ø* or an A-nucleus as in *-dAn* or *-lAr*, or a combination of the two as in *-lArØ*. While root vowels are fixed (i.e., the harmonic form is stable), suffix vowels alternate depending on the element content of the final stem vowel.

(87) *Harmony in Turkish suffixes*

Nom sg	Poss	Abl	Nom pl	Poss/Acc pl	
	-(s)Ø	*-dAn*	*-lAr*	*-lArØ*	
oda	odasɨ	odadan	odalar	odalarɨ	'room'
son	sonu	sondan	sonlar	sonlarɨ	'end'
boru	borusu	borudan	borular	borularɨ	'pipe'
köy	köyü	köyden	köyler	köyleri	'village'
kurt	kurdu	kurttan	kurtlar	kurtlarɨ	'worm'
tilki	tilkisi	tilkiden	tilkiler	tilkileri	'fox'
inek	inei	inekten	inekler	inekleri	'cow'
dere	deresi	dereden	dereler	dereleri	'river'

at	atɨ	attan	atlar	atlarɨ	'horse'
tat	tadɨ	tattan	tatlar	tatlarɨ	'taste'
kɨz	kɨzɨ	kɨzdan	kɨzlar	kɨzlarɨ	'girl'
kap	kabɨ	kaptan	kaplar	kaplarɨ	'container'
ütü	ütüsü	ütüden	ütüler	ütüleri	'iron'

When the context for harmony is missing, the empty nucleus in a suffix is realized as [ɨ] and the A-nucleus as [a]. Otherwise, harmony generates [i, u, ü] in empty nuclei, and [e] in A-nuclei. The empty nucleus in *–lArØ* only ever alternates between [i] and [ɨ], because the preceding A-nucleus is neither a target of nor transparent to U-harmony. Similarly, in the form [odasɨ], the final empty nucleus is not realized as [u] despite the presence of the element U in the initial nucleus. The intervening A-nucleus does not undergo U-harmony, and it blocks the progression of U-harmony.

To summarize, in typical roots and suffixes, all nuclei but the first will either be lexically empty or contain only the A element. The surface form will depend on the first nucleus, which may contain any one of the seven true vowels of Turkish (default [ɨ] is not present in underlying forms). Consequently, at the surface, non-initial nuclei are normally only realized as [ɨ, a, i, u, e, ü]. The vowels [o, ö] can only occur in initial nuclei of harmonic forms. With the element content U-A and U-A-I respectively, these vowels cannot arise as a result of harmony: if the underlying nucleus is empty, it cannot receive a harmonic A element; if the underlying nucleus contains the A element, it cannot be a target of U-harmony.

Clements and Sezer (1982), however, present a number of disharmonic exceptions. These are forms with [o, ö] in non-initial nuclei, or else they are instances of [i, u, e, ü] that cannot be attributed to spreading (88).

(88) *Disharmonic forms*

tak<u>o</u>z 'wooden wedge'

šoför	'driver'
takvim	'calender'
marul	'lettuce'
haber	'news'
virüs	'virus'

These forms exceptionally contain non-initial nuclei that are lexically specified for more than just the A-element or an empty nucleus. If non-initial vowels can be specified, this of course raises the question of whether the vowels in apparently harmonic roots are not also lexically specified, rather than being actively derived via spreading. Unlike suffix vowels, root vowels do not alternate, so it could be that root harmony is merely apparent. This is the conclusion that Clements and Sezer arrive at, but evidence from a language game suggests their conclusion is misguided.

Harrison and Kaun (2001) taught some Turkish speakers a language game involving whole-word reduplication, along with replacement of the initial vowel by [a]. With [a] in initial position, any context for harmony is removed, since the harmonic elements U and I are absent. Of course, if harmony is not a synchronic process, the removal of apparent triggers should have no effect. But this is not what we find (89).

(89) *Reduplication language game*

Input	Reduplicant	
a. kibrit	kabrɨt	'match'
b. bütün	batɨn	'whole'
c. butik	batik	'boutique'

When the input is a harmonic root (89a,b), the second vowel of the reduplicant is affected by removal of the harmonic source. In this case, the vowel can only be default [ɨ] (or hypothetically [a] in the case of an A-nucleus – no examples are given in the study). When the input is a disharmonic root (89c), on the other hand, the second vowel is unaffected. The simplest way to account for this divergence in behaviour is to attribute the quality of the input second vowel in (89a,b) to spreading from the initial vowel. Thus, only the second vowel in the disharmonic input in (89c) is lexically specified. Importantly, this finding supports my view that the second nucleus in [kibrit] and [bütün] is underlyingly empty, receiving its surface content via spreading of I or I-U respectively from the initial nucleus. It also confirms my contention that [ɨ] is simply a default vowel that appears when harmony cannot apply.

This is not quite the whole story, however. There are in fact two types of empty nuclei in the data: those that participate in harmony and those that remain empty. Both types of empty nucleus can appear in either final or medial position. Some examples of the two types of final empty nuclei are given below.

(90) *Two types of final empty nuclei in Turkish*

a. <u>With default/harmonic vowel</u>		b. <u>With Ø</u>	
karɨ	'wife'	karØ	'snow'
kadɨ	'judge'	tatØ	'taste'
valɨ	'villa'	markulØ	'lettuce'
kuru	'dry'	ugurØ	'fortune'
oku	'read'	nutukØ	'speech'
kiši	'person'	buzØ	'ice'

To account for the divergent behaviour of final empty nuclei, Charette (2006) adopts the same solution that Brockhaus (1995) advocated for similar cases in German. That is, Charette proposes that the final empty nucleus parameter is on in Turkish. Consequently, empty nuclei may remain empty: [karØ]. When they do not remain empty, however, this is because there is an additional empty CØ syllable: [karɨCØ]. In this case, the final empty nucleus is licensed to remain empty, but the preceding nucleus fails to be properly governed and must thus be realized. The same criticisms that were raised in discussing the German data apply here, including the issue of the learnability of such an abstract form.

In addition, Charette's account fails to address the question of medial empty nuclei that either do or do not remain empty. Most of the time, medial empty nuclei do not remain empty. In examples such as [bütün], it can be argued that the second nucleus cannot remain empty because it is not properly governed. Crucial evidence will come therefore from: i) empty nuclei that do not remain empty even when they are potentially properly governed; and ii) empty nuclei that are not properly governed but nonetheless remain empty. In order to determine where an unrealized empty nucleus should be posited, however, it is necessary to review the syllable structure of Turkish.

Turkish is usually claimed to have (C)V(C) syllable structure: rhymes may branch, but onsets do not. The view that Turkish lacks branching onsets is supported by adaptations of loanwords with initial branching onsets: [tɨrafik] 'traffic', [buluǰin] 'blue jeans' (Clements & Sezer 1982).[68] Codas may result from gemination. Otherwise, the set of coda consonants is somewhat difficult to determine. Still, Clements and Sezer identify the following final consonant sequences as constituting coda-onset pairs.[69]

(91) *Final consonant sequences*

Sonorant + obstruent:	kent	'city'	harf	'letter'
Voiceless fricative + stop:	čift	'couple'		
[ks]:	raks	'dance'		

[68] These adaptations are in fact optional, in the sense that the consonant sequences [tr] and [bl] may be realized without an epenthetic vowel in careful speech. In the usual colloquial form, nonetheless, the cluster is broken up as shown.

[69] Harris (1997) makes the (frankly more plausible) claim that Turkish codas are limited to the universally preferred set [r, l, s, N]. However, I have not found independent confirmation of this claim. For my purposes, nothing essential hangs on this issue.

Based on these examples, we must assume that Turkish is quite permissive in what appears in a coda. All the same, Turkish codas must respect universal principles. For example, coda nasals should be homorganic with an adjacent onset. Likewise, coda-onset sequences should conform to the complexity condition.

Based on these principles, and keeping in mind that onsets do not branch, we can identify various forms that simply must contain an unrealized empty nucleus.

(92) *Medial empty nuclei that remain empty*

takvim	=	takØvim	'calender'
gɨrtlak	=	gɨrtØlak	'throat'
damɨtmak	=	damɨtØmak	'distill'
kibrit	=	kibØrit	'match'
kaplɨja	=	kapØlɨja	'thermal spring'

Admittedly, all of the empty nuclei posited in (92) are followed by a realized nucleus and hence potentially properly governed. In this case, it could be unnecessary to posit the existence of two types of empty nuclei in Turkish. The problem is that potentially properly governed empty nuclei do not always remain empty. In the form [kapØlɨja], the penultimate empty nucleus with default [ɨ] is potentially properly governed by the final nucleus. Nonetheless it is realized and apparently properly governs the preceding empty nucleus. This leads to the output as it stands rather than to the output that would be expected if proper government operates right-to-left: *[kapɨlØja].

Hypothetically, of course, proper government might operate left-to-right in Turkish. This solution is undermined by [damɨtØmakØ], which presents the reverse situation of [kapØlɨja]. In [damɨtØmakØ], the antepenultimate nucleus remains empty and the preceding nucleus contains default [ɨ]. Forms such as [rabɨta] 'connection' compound the situation, since the default [ɨ] has a potential proper governor on either side.

The better solution is to countenance two types of empty nuclei: those with a root node, which are realized either by harmonic vowels or by default [ɨ]; and those without a root node, which remain empty. The alternative solution (considered previously here) is to treat nuclei realized as [ɨ] as in fact filled nuclei, lexically specified for the elements @-I (i.e., following the element content I proposed in Chapter II). In this case, only unrealized nuclei would be empty. One argument in favour of this competing view is that [ɨ] may be stressed in Turkish. Typically stress falls on the last realized nucleus in a word, and if this final nucleus happens to contain [ɨ], it will be stressed. This contrasts with the unstressability of schwa in French, German, English, and Palauan.

Remember that in French, when an empty nucleus would appear in a stressed position, rather than harbouring default schwa, it is realized as [ɛ]. In German and English, schwa is simply never stressed. The inability of schwa to support stress was used earlier as an argument in support of viewing schwa nuclei as in fact empty. But it does not necessarily follow from the stress-bearing capacity of Turkish [ɨ] that it must be a filled nucleus in this language. It is always possible that, in French and German, stress is lexically specified or else determined early in a derivation, before the empty nucleus receives any content. Conversely, stress in Turkish might be determined after the nucleus has been assigned default (or harmonic) content. The presence of content in the lexically empty nucleus may then permit it to bear stress. Alternately, there may be something special about [ɨ] that allows it to bear stress; perhaps its having (as I proposed) both I and @ as content counts in its favour, whereas the fact that [ə] has only @ somehow limits its stress-bearing capacity.

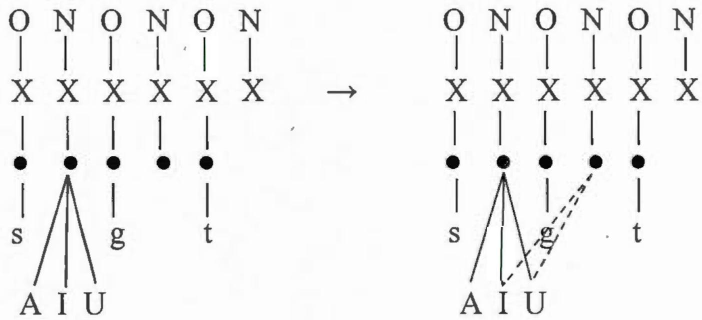
There are other arguments against the view that [ɨ] represents a filled nucleus. First, if [ɨ] contains the elements @-I and Turkish has I-harmony, one wonders why [ɨ] itself never triggers I-harmony. For example, in [kɨyak] 'excellent', the second nucleus contains the element A, making it a potential target for I-harmony, yet I fails to spread from [ɨ]. If [ɨ] is lexically specified for I-@, this situation is hard to explain. If, on the other hand, I is simply part of the default realization of the empty nucleus, it may be inserted after harmony has applied.

Next, when an [ɨ]-nucleus is itself the target of harmony, neither of the elements @ or I contribute to the surface form. For example, the suffix *-(s)Ø* is realized with [ɨ] in [odasɨ], but with [u] in [borusu]. If the suffix nucleus contains a fully specified vowel [ɨ], what happens when U-harmony applies? Why isn't the nucleus realized as [ü] (I-U) or even as [Ü] (I-U-@)? Under the filled-nucleus analysis, the surface form should be *[borusü] or *[borusÜ]. Under the empty nucleus analysis, the nucleus can be realized either with default [ɨ] or with a harmonic vowel. The default content I-@ only appears if harmony fails to apply. This explains the absence of a combination of I-@ and harmonic elements.

In sum, under my analysis, a form such as [damɨtØmakØ] contains two types of empty nuclei with different degrees of degeneracy (93).

(93)
```
O N O N O N O N O N        O N O N O N O N O N
| | | | | | | | | |        | | | | | | | | | |
X X X X X X X X X X   →    X X X X X X X X X X
| | | | |   | | |          | | | | |   | | |
• • • • •   • • •          • • • • •   • • •
| | |   |   | | |          | | | : |   | | |
d a m   t   m a k          d a m ɨ t   m a k
```

Distinguishing between the empty nuclei in [damɨtØmakØ] based on the presence versus absence of a root node is motivated by the notion of autosegmental licensing. Only the form containing a root node is able to license default content such as [ɨ]. In addition, the notion that it is specifically the presence versus absence of a root node that distinguishes the two types of empty nuclei is supported by the facts of harmony. This can be shown with a form such as [sögüt] where both I- and U-harmony apply.

(94) 

Only the elements I and U spread from the first to the second nucleus in [sögüt]. The element A does not participate in harmony. If it were simply a question of the entire segment spreading from the first to the second nucleus, the presence of a root node would be obviated in the target nucleus. The root node for the first vowel could simply be doubly associated to two nuclei (as shown previously with respect to empty nuclei in Selayarese and Winnebago). However, the fact that only certain elements in the vowel spread means that a receptive root node must be present in the target. Indeed, the absence of a root node in the third nucleus is what excludes it from participating in this form of harmony. The representational distinction I propose for the two types of empty nuclei is thus robustly supported by their behaviour in Turkish.

4.4 Empty nuclei in initial position

Empty nuclei in initial position were first proposed by Kaye (1992) to account for sC-clusters, which have certain special properties. Generally a sequence of initial consonants is assumed to form a branching onset. However, sC-clusters show properties that are inconsistent with a branching-onset analysis. First, they do not have the falling complexity differential expected of branching onsets. Next, they differ from true branching onsets in how they participate in certain morphophonological processes. Consequently, Kaye proposed to analyze [sC] as a coda-onset sequence. For licensing purposes, an [sC] coda-onset at the beginning of words is necessarily preceded by an empty nucleus (the empty nucleus licenses the dependent coda position in the rhyme). Other initial configurations that can be analyzed as coda-onset sequences preceded by an empty nucleus are full and partial geminates (Piggott 2003b). These also fail to exhibit the profile of branching onsets, and they behave phonologically like coda-onset sequences. The evidence for analyzing initial sC-clusters and full and partial geminates as coda-onsets with a preceding empty nucleus will be reviewed in turn below. Particular emphasis is placed on the evidence for initial full and partial geminates, since these have not traditionally been covered in the GP literature.

After this, I will address the question of whether initial empty nuclei occur both with and without a root node. Intriguingly, there seems to be a decided preference for such nuclei to appear without a root node and to remain empty. Indeed, I have uncovered no cases of initial empty nuclei (either with or without a root node) that are targets of harmony. The instances of initial empty nuclei with a default vowel in the unspecified root node are marginal. Why this should be remains a mystery. Apparently, just as final position shows a preference for empty nuclei without a root node (hence the universal tendency for apocope), initial position also favours the more degenerate form of empty nucleus. This is particularly surprising given that the initial empty nucleus heads a rhyme that licenses a dependent coda. A priori, one would expect an empty nucleus with a root node to be better equipped to fill this role. A thorough explanation of this preference for more degenerate empty nuclei at the left edge remains elusive, but provisionally we might attribute this phenomenon to a demarcatory function.

4.4.1 Empty nuclei before initial sC-clusters

The difference between true branching onsets and sC-clusters can be demonstrated with reference to English. In English, fricatives such as [f, θ] can combine only with liquids [fl, θr] or glides [fj, θw] at the beginning of words. These combinations have a falling complexity profile, which is consistent with a branching-onset analysis. On the other hand, [s] can combine with stops [sp, st, sk] and with nasals [sn, sm], both of which contravene the complexity condition applying to branching onsets: [s] (with elements h-R) is less complex than stops or nasals (with three elements).[70] In addition, [s] can appear at the beginning of three-member clusters such as [spl, str, skw]. If constituents are maximally binary branching, these sequences are doubly illicit as branching onsets.

Further problems for a branching-onset analysis of sC-clusters comes from how [s] combines with liquids. Other coronal obstruents such as [t, θ] combine with [r] but not with [l] (presumably because [tl] and [θl] constitute an OCP violation[71]). Examples of [tr] and [θr] abound: *tree, true, tread,* and *three, through, thread.* Words beginning with [tl] and [θl], on the other hand, are non-existant, and in the few instances where [tl] and [θl] sequences appear within words (e.g., *atlas, athlete*), these are heterosyllabic sequences, not branching onsets. Remarkably, the reverse situation is found with sC-clusters: [s] combines freely with [l] (see *sleep, slow, slim*) but not with [r]. In the rare instances where a [sr] sequence might be posited, as in *Sri Lanka* or *Sree Sreenivasan,* this is usually repaired to [šr]. Interestingly, [š] is like [t] and [θ] in English: while it combines freely with [r] (e.g., *shriek, shrew, shrimp*), it does not normally combine with [l], except in Yiddish loanwords (e.g., *schlep, schlemiel, schlock*). The contrasting distribution of [s] and other voiceless coronal obstruents with liquids is presented below in (95); for comparison, the ability of all these coronal obstruents to combine with the glide [w] is also given.

---

[70] In languages such as German and Yiddish, šC-clusters are found instead of sC-clusters. See, for example, German *Strasse* 'street', *spucken* 'to spit', *Schnee* 'snow', and *schlafen* 'to sleep', all with initial [š], or Yiddish loanwords to English *such as schtick, spiel, schnoz* and *schlep.* Their behaviour parallels that of English sC-clusters.

[71] According to Harris (1994), the members of a branching onset may have maximally two elements in common. Since [t] or [d] and [l] share both R and ?, these are ruled out as branching onsets. This appears not to be the whole picture, however, since [š] and [r] can co-occur in a branching onset, yet these share two elements also, namely R and I.

(95) *Initial voiceless coronal obstruent + liquid/glide sequences in English*

Rhotic liquid	Lateral liquid	Glide
[tr] – *tree, true, tread*	*[tl]	[tw] – *twin*
[θr] – *three, through, thread*	*[θl]	[θw] – *thwart*
[šr] – *shriek, shred, shrimp*	*[šl]? – (*schlep*)	[šw] – *schwa*
*[sr]	[sl] – *sleep, slow, slim*	[sw] – *sweat*

In essence, [s] repeatedly behaves differently in terms of the consonants it combines with: it combines with stops and nasals; it forms three-member clusters; and [s] can be followed by [l], but not [r]. If these are branching onsets, their profile is unlike that of any other branching onsets. The only combination in which an sC-cluster potentially forms a true branching onset in English is [sw].

If sC-clusters are not branching onsets, then how are they to be analyzed? Kaye (1992) argues that sC-clusters are coda-onset sequences preceded by an empty nucleus, as shown in Figure 4.4.

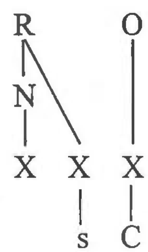


Figure 4.4 *Representation of initial sC-clusters*[72]

Support for this analysis of sC-clusters comes from their exceptional behaviour in phonotactics and phonological processes. For example, in Italian, stressed syllables require a heavy rhyme, comprised of either a long vowel or a short vowel followed by

[72] See Nikièma (2003) for a competing analysis of sC-clusters as a series of two onsets separated by an empty nucleus. Without going into details, various inconsistencies make this analysis less convincing than the coda-onset analysis, however, so I will not examine it in depth here.

a coda. Consequently, stressed vowels are long before both singleton and branching onsets (see [fá:to] 'fate' and [ká:pra] 'goat'), but short before coda-onset sequences (see [fátto] 'fact' or [párko] 'park'). Before sC-clusters, stressed vowels are also short (see [pésta] 'trail' or [móska] 'fly'), a fact that is consistent with their being coda-onsets.

If word-internal sC-clusters are coda-onset sequences, the same analysis should also apply to initial sC-clusters. Kaye presents evidence from Italian, Ancient Greek, European Portuguese, and British English to support this conclusion (see Goad 2012 for further evidence). Here I review only the evidence provided by *raddoppiamento sintattico* in Italian. This sandhi process involves the gemination of a word-initial consonant when the preceding word ends in a short stressed vowel (Nespor & Vogel 1979; Chierchia 1986). The process provides the final light stressed syllable with a coda, thus satisfying the requirement that stressed rhymes be heavy. As shown below in (96), the process applies whether the second word begins with a singleton or branching onset; it does not apply, however, before sC-clusters.

(96) Raddoppiamento sintattico *before:*

a. Singleton onset

paltó pulíto	[paltó**pp**ulíto]	'clean coat'
é caríno	[é**kk**aríno]	'it is pretty'

b. Branching onset

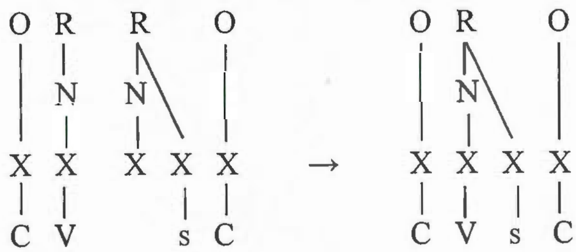
cittá tríste	[čittá**tt**ríste]	'sad city'
caffé fréddo	[kaffé**ff**réddo]	'iced coffee'

c. sC-cluster

cittá straniéra	[čittá**s**traniéra] / *[**ss**]	'foreign city'
caffé spésso	[kaffé**s**pésso] / *[**ss**]	'thick coffee'

If sC-clusters are coda-onsets, the failure of *raddoppiamento sintattico* to apply has a simple explanation: there is no need for [s] to geminate since it can be parsed into the coda of the preceding syllable as is. This process is illustrated in (97).

(97) *No* raddoppiamento sintattico *before sC-clusters*



caffé spésso [kafféspésso]

When a word ending in a stressed vowel precedes a form beginning with an sC-cluster, the empty nucleus fuses with the preceding filled nucleus, and the [s] syllabifies into the rhyme headed by the filled nucleus.

In languages like Italian and English, sC-clusters are the only kind of coda-onset sequence found after initial empty nuclei. In other languages, however, a case can be made for treating initial full and partial geminates as coda-onset sequences preceded by an empty nucleus, a point first made by Piggott (2003b) regarding Selayarese.

4.4.2 Empty nuclei before initial full and partial geminates

Some languages allow initial sequences of two identical consonants or of a homorganic nasal and obstruent: CC or NC. These sequences are particularly common in Austronesian and Bantu languages. Such full and partial geminates, like initial sC-clusters, must be syllabified as coda-onsets preceded by empty nuclei. Examples from Selayarese (Mithun & Basri 1986; Piggott 2003b) and LuGanda (Clements 1986) are provided below in (98).

(98) *Initial full and partial geminates*

Selayarese (Autronesian)		LuGanda (Bantu)	
a. Initial CC:		a. Initial CC:	
[ppalui]	'he/she is cooking'	[ttama]	'cheek'
[mmari]	'to stop'	[ggwowo]	'savoury smell'
b. Initial NC:		b. Initial NC:	
[ntale]	'to cross'	[mpisi]	'hyena'
[ŋgora]	'to shout, cheer'	[mbwa]	'dog'

Like sC-clusters, none of these initial sequences are consistent with a branching-onset analysis. In terms of element structure, they fail to instantiate a falling complexity cline: CC and NC sequences have segments with the same number of elements (three). Moreover, since Selayarese disallows sequences of consonants with a typical branching-onset profile (e.g., C + liquid or glide), branching onsets appear to be illicit. LuGanda does allow branching onsets, but only C + glide. In addition, as shown above in [ggwowo] and [mbwa], glides in LuGanda may appear *after* a CC or NC sequence. Like sCC-clusters, the [ggw]/[mbw] sequences appearing in (48) cannot be onsets, since branching constituents are maximally binary. Finally, word-internally, the same CC and NC sequences are the only attested coda-onset pairs (99).[73]

(99) *Initial full and partial geminates*

Selayarese		LuGanda	
a. Medial CC:		a. Medial CC:	
[ruppa]	'face'	[maŋŋo]	'teeth'
[bannaŋ]	'thread'	[maggwa]	'thorns'

[73] Selayarese also has initial (and medial) ʔC sequences. These will be addressed later in Chapter V, which deals with empty codas.

b. Medial NC:		b. Medial NC:	
[miŋka]	'but'	[ku:mpi]	'near'
[ʔondaŋ]	'to invite'	[endwadde]	'disease'

The implication is that both initial and medial CC and NC sequences should be analyzed as coda-onset sequences. The only real difference is that the initial full and partial geminates are preceded by an empty as opposed to a filled nucleus.[74]

The evidence from Selayarese and LuGanda is quite clear: the combinations of segments forming initial CC and NC sequences are consistent with their being coda-onset sequences rather than branching onsets. Further evidence for this analysis comes from phonological processes involving full geminates in Trukese (Austronesian) and partial geminates in numerous Bantu languages.

Geminates abound in Trukese (Hart 1991): all consonants in the inventory geminate, and geminates may be medial or initial. Indeed, the only codas found in Trukese are formed via gemination. Interestingly, the presence of an initial geminate in a noun influences the way the noun behaves with respect to certain phonological processes.

To explain, Trukese nouns are subject to a minimal word constraint: they must contain at least two syllables or, in monosyllables, a long vowel (but cf. Hart 1991 for a moraic analysis). A minimal noun is thus CVCV or CVV (i.e., CV:). In addition, final long vowels undergo shortening, and final short vowels delete (i.e., in either case, the final timing slot is severed). Examples (with Trukese orthography) are given in (100).

[74] An issue that is addressed later on in this section is that NC sequences in Bantu languages such as LuGanda do not necessarily surface as coda-onsets, although I argue that this is still the underlying syllabic parse. Instead, NC may surface as a prenasalized stop or else the nasal may even appear in the initial empty nucleus (e.g., when the nasal constitutes a tone- or stress-bearing unit). In Selayarese, on the other hand, true NC sequences always surface as coda-onsets, and word-medially, there is in fact a contrast between NC and prenasalized ^{N}C (e.g., [bombaŋ] 'wave' vs. [bo:^{m}baŋ] 'bamboo skin').

(100) *Shortening/deletion of final vowels*

a. V: → V

pechee	→	peche	'foot'
tikkaa	→	tikka	'coconut oil'
chuuchuu	→	chuuchu	'urine'

b. V → Ø

omosu	→	omos	'turban shell'
nemeneme	→	nemenem	'authority'
mékúré[75]	→	mékúr	'head'

Note that the output of the process in each of the forms above is always a minimal word. When the input is a monosyllabic CVV form, however, the application of vowel shortening would produce illicit CV output. Consequently, vowel shortening is blocked in monosyllabic forms, and the final timing slot is maintained (101a). Nonetheless, vowel shortening does apply to apparent monosyllables when they contain an initial geminate (101b).

(101) a. *No vowel shortening in monosyllables:*

téé	→	téé	'islet'
chúú	→	chúú	'bone'
máá	→	máá	'death'

b. *Vowel shortening in apparent monosyllables with initial geminate:*

ssóó	→	ssó	'thwart of canoe'
kkáá	→	kká	'taro sp.'
ttoo	→	tto	'cry'

---

[75] An accent over a vowel in the Trukese data is an orthographic convention to distinguish segments. For example, 'e' is [e], but 'é' is [ʌ]. See Hart (1991) for details.

The implication here is that such apparent monosyllables are in fact bisyllabic. The initial geminate constitutes a coda-onset sequence preceded by an empty nucleus. Consequently, the vowel is able to shorten without incurring a violation of the minimal word constraint.

In a further twist, final vowel deletion applies to bisyllabic forms, but the preceding vowel undergoes lengthening so as to conform to the minimal word requirement.

(102) *Vowel deletion and compensatory lengthening in bisyllables:*

wutu	→	wuut	'interior'
tipe	→	tiip	'emotions'
chúká	→	chúúk	'basket'

This lengthening applies only in bisyllabic forms, which require the long vowel in order for the surface form to constitute a minimal word. When the final vowel deletes from polysyllabic forms such as [omosu] → [omos] in (100b) above, there is no concomitant vowel lengthening.

Interestingly, vowel lengthening also fails to apply in apparent bisyllabic forms with an initial geminate (103).

(103) *No compensatory lengthening in bisyllables with initial geminate:*

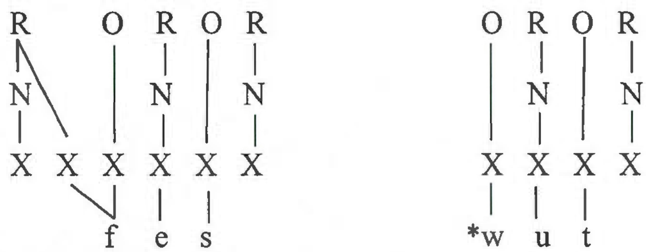
ffesi	→	ffes	'pool of money'
ttina	→	ttin	'shine'
nnifi	→	nnif	'wooden mortar'

Again the implication is that vowel lengthening is obviated in these forms by the presence of an initial empty nucleus. For the purposes of determining the minimal word, then, the initial syllable clearly counts. This is doubly interesting because the final empty nucleus that results from vowel deletion in the form *wutu* → **wut* does not contribute to the minimal word (hence the application of vowel lengthening:

wuut). In a nutshell, the syllable headed by initial but not final empty nuclei contributes to meeting the minimal word requirement.

The obvious distinction between the initial syllable in *ffes* and the final syllable in **wut* is that the rhyme of the former is not completely empty, as shown in (104).

(104) a. <u>Output *ffes*</u> b. <u>Output **wut*</u>



The ability of the initial empty nucleus to contribute to the minimal word can thus be attributed to its having a filled coda. A coda formed by gemination does not contain any elements of its own. Nonetheless, by virtue of the content spreading from the neighbouring onset, the geminate coda establishes the weight of the initial rhyme, permitting it to contribute to the minimal word. As we saw previously with the Italian phenomenon of *raddoppiamento sintattico*, a geminate coda can count in the calculation of weight. In sum, the evidence from Trukese supports an analysis of initial geminates as coda-onset sequences preceded by empty nuclei. Additional evidence involving initial partial geminates appears in various Bantu languages.

The evidence from Trukese suggests clearly that its initial full geminates constitute a coda-onset pair with a preceding empty nucleus. In languages with homorganic NC sequences, however, the situation is not always so clear-cut. Particularly in initial position, it can be difficult to determine whether NC constitutes two discrete segments forming a coda-onset pair or a single complex segment syllabified into the onset. Odden (in prep.) describes the Bantu language Kikerewe as being a no-coda language with only open syllables. His assertion rests on the assumption that initial and medial NC sequences in Kikerewe form contour segments in onset position. Indeed this is probably the correct analysis of the surface form in Kikerewe and other Bantu languages. Nonetheless there is reason to believe that underlyingly NC sequences are parsed into separate coda and onset constituents.

To demonstrate, a process of vowel lengthening commonly applies before NC sequences in Bantu languages. This process implies a derivation whereby NC starts as a heterosyllabic pair of segments, but the nasal subsequently associates to the onset timing slot, leaving its own slot available for the preceding vowel. The lengthening process can be illustrated by data from Jita (Downing 2005). This language has both long and short vowels before single consonants, but only long vowels before NC.

(105) *Distribution of long and short vowels in Jita*

Long vowels		Short vowels	
a. *Before single consonants:*			
oku-**cuu**ma	'to jump'	oku-c**u**ma	'to get wealthy'
oku-β**ii**ka	'to keep safely'	oku-β**i**ka	'to announce a death'
b. *Before NC:*			
oku-**fuu**mbula	'to guess a riddle'	XX	
oku-ki**i**ŋga	'to arrive'	XX	

If NC sequences are coda-onset segment pairs, one might expect the exact reverse distribution of vowel length before NC: only short vowels should be attested. On the other hand, if NC sequences form a complex onset, one would expect both long and short vowels to be able to precede them. In other words, the fact that only long vowels are found before NC sequences appears consistent with neither analysis.

As shown below in Figure 4.5, the solution is to treat NC sequences as complex onsets on the surface, but as derived from underlying coda-onset sequences. This analysis was originally proposed by Clements (1986) for a similar situation in LuGanda.[76]

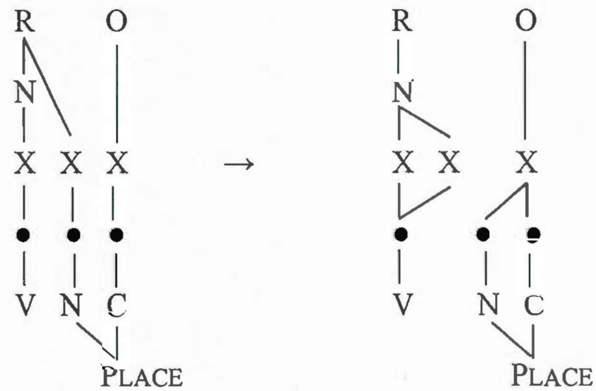


Figure 4.5 *Compensatory lengthening before NC sequences*

Given the process of compensatory lengthening, necessarily vowels before NC sequences will be long since, in re-associating to the onset position, the nasal leaves its timing slot available for the preceding vowel. This analysis thus explains the absence of short vowels before NC sequences.

Generally in Bantu languages, the same process of vowel lengthening applies to a CV prefix attached to an NC-initial stem. The implication is that, like medial NC, the initial NC is lexically represented as a coda-onset sequence, but the nasal undergoes a process of re-association. The following examples from Kikerewe (Odden 1995) show that the prefixes *mu-* and *ha-* lengthen before NC-initial stems (106a), but not before C-initial stems (106b).[77]

---

[76] See Downing (2005) for an alternative representation using moras. In her view, the nasal remains syllabified as a coda, and the vowel lengthens by also associating to the mora dominating the nasal, with the result that this mora is doubly linked to the vowel and nasal.

[77] Kikerewe contains an unusual exception: although vowel lengthening applies to CV-prefixes such as *mu-* and *ha-*, it does not apply to V-prefixes such as *e-* (e.g., *e-ndálá* 'leopard', *e-mbúzi* 'goat', and *e-ŋgezi* 'flood'). Apparently, for reasons that remain obscure, the nasal does not re-associate to the adjacent onset following a V-prefix. Hence, no timing slot is available to feed lengthening.

(106) *Kikerewe V-lengthening in CV-prefixes*

a. Before NC-initial stems:		b. Before C-initial stems:	
muu-ndálá	'in the leopard'	mu-chúúmba	'in the room'
haa-mbúzi	'at the goat'	ha-chúúmba	'at the room'
muu-mpéte	'in the ring'	mu-mabáále	'in the stones'
haa-ndíílo	'at the relish'	ha-mabáále	'at the stones'

Initial NC sequences in Bantu languages do not necessarily become prenasalized stops when not preceded by a prefix. Instead, the nasal appears to re-associate to the initial empty nucleus.[78] Evidence comes from stress placement in Swahili and from tones in LuGanda. Regular stress falls on the penultimate syllable in Swahili (Brandon 1975, cited in Downing 2005). When a form consists of an NC sequence followed by only a single vowel, stress (indicated by underlining) falls on the nasal.

(107) *Swahili nasal class 9 prefix*

n-zuri	'good'
m-baya	'bad'
m-pya	'new'

Further evidence comes from tones in LuGanda (Herbert 1975). In this language, the nasal in an initial NC sequence functions as a tone-bearing unit. Again, this behaviour suggests that it cannot be part of the onset; instead it appears to constitute a separate syllable.

(108) *Tone-bearing initial nasals in LuGanda*

ɲ-ju-ki	ɲ̀júkì	'(the) bee'
m-pa	m̀pâ	'give me'
ŋ-ku-ba	ŋ̀kúbà	'rain'

[78] I am grateful to Mohamed Guerssel (personal communication) for this insight.

In brief, when the nasal in an NC sequence is preceded by a vowel in Bantu languages, typically the nasal re-associates with the onset C to form a contour segment. When there is no preceding vowel, on the other hand, the nasal remains in the initial syllable, re-associating to the empty nucleus. By occupying the nucleus of this initial syllable, the nasal can attract both stress and tone.

In sum, evidence from various languages, both in terms of the combinations of segments involved and in terms phonological processes, points to the conclusion that certain initial sequences of consonants form coda-onsets. These include sC-clusters and full and partial geminates.[79] For licensing purposes, these coda-onsets must be preceded by empty nuclei.

From a GP standpoint, the presence of an empty nucleus before an initial coda-onset sequence is problematic since this nucleus cannot be licensed by proper government. Proper government is blocked because a governing domain (the coda-onset) comes between the two nuclei. Consequently, Kaye (1992) proposed provisionally that sC-clusters have a special property that allows the empty nucleus to be licensed by 'magic'. Since 'magic' licensing is parametric, it does not apply in all languages. For languages such as English, Italian and Ancient Greek that permit initial sC-clusters, the parameter is on; for Spanish, Brazilian Portuguese and Farsi, it is off. Conceivably, this special licensing ability could be extended to include empty nuclei before CC and NC sequences. In the next section, however, we examine another way of capturing empty nuclei that remain empty or, alternately, that are expressed, and that is to distinguish between empty nuclei with or without a root node.

[79] Note that I do not entertain the possibility of parsing initial CC and NC sequences as a pair of onsets separated by an empty nucleus. As indicated in previous note 72, this analysis has been proposed by Nikièma (2003) to handle sC-clusters, but to my mind the arguments he marshals fail to convince. For CC and NC sequences, the analysis strikes me as particularly ill-conceived given that these sequences are canonical coda-onsets word-internally.

4.4.3 On the representation of initial empty nuclei

Having established the existence of initial empty nuclei, we can now consider the nature of their representation. Conventionally, GP represents an initial empty nucleus before an sC-cluster with a timing slot as shown in Figure 4.6.[80] The nucleus is required in order to license the coda dependent, and if, as GP assumes, licensing relations are established between positions on the timing tier, naturally the empty nucleus must have a timing slot. By extension, empty nuclei before partial and full geminates should also be represented with a timing slot.

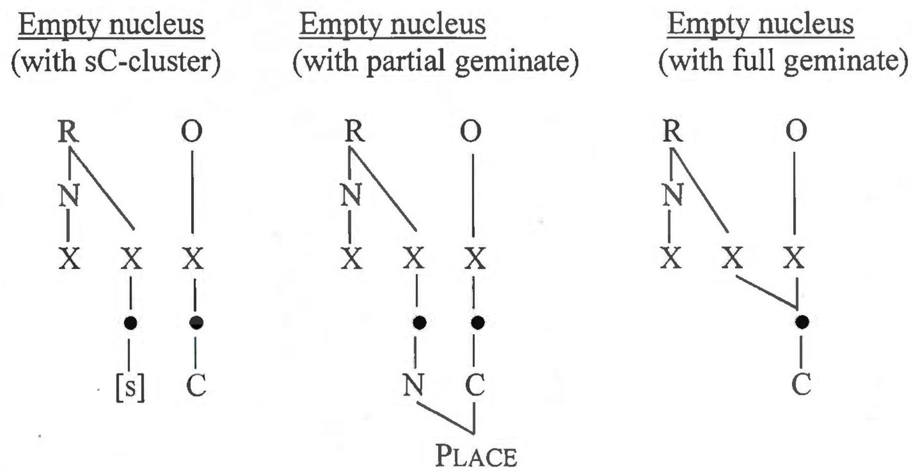


Figure 4.6 *Initial empty nuclei*

In languages which do not permit these configurations with initial empty nuclei, the structure must be repaired. This can be achieved notably by realizing a default vowel in the empty nucleus. For example, words with initial NC sequences produced by

[80] Technically, Kaye (1992) also includes an empty onset before the empty nucleus. This is omitted here partly to keep the representation as uncluttered as possible. Also, whether an empty onset actually figures in such contexts is open to debate.

English speakers are generally rendered with an initial schwa as shown in the examples here.[81]

(109) *English pronunciation of NC sequences*

[ə]*Mbuti*	'a Pygmy people of western Uganda'
[ə]*Ngorongoro*	'a crater in Tanzania'
[ə]*Ndola*	'a city in Zambia'

Similarly, sC-clusters in loanwords to Turkish are modified by adding an initial vowel (van der Hulst & van de Weijer 1991).[82] Typically, this is the default vowel [ɨ] that appears in empty Ns with a root node – exceptions such as [iskelet] with initial [i] follow the pattern seen previously for harmonic forms.

(110) *Turkish loanword adaptation of sC-clusters*

ɨspanak	'spinach'
ɨstatistik	'statistics'
iskelet	'skeleton'

The question, of course, is whether the representation of these forms is the same in languages where the empty nucleus remains unexpressed and in those where it receives default expression. The standard GP position is that the representation is the same, but that in some languages the initial empty nucleus is licensed to remain empty, while in others the empty nucleus is unlicensed and must hence be expressed.

The only truly problematic situation for the GP position would come from languages that permit both. That is, the language might have minimal pairs such as [mbuti] and

---

[81] For confirmation, see the pronunciation with initial schwa given for *Mbuti* and *Ngorongoro* at http://oxforddictionaries.com. Arguably, however, a pronunciation with the initial full vowel [ɛ] is also possible, in which case a filled nucleus is inserted to repair the structure.

[82] Van der Hulst and van de Weijer also give [sɨmokin] for 'dinner jacket' (from French *le smoking*), so vowel insertion in initial position may require that the consonant following [s] be an oral stop.

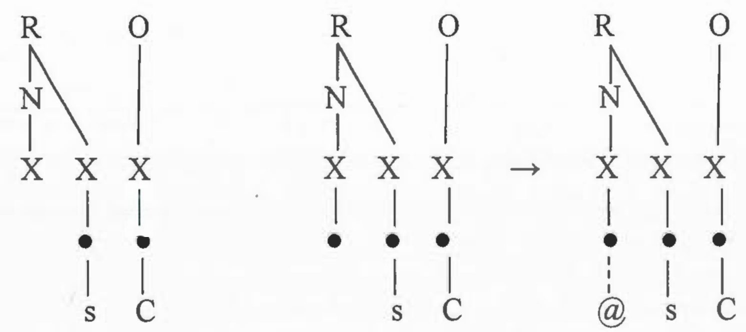
[əmbuti], [ndola] and [əndola]. Assuming the schwa in such forms is truly a default expression of an empty nucleus, there must be some difference between the empty nuclei at the beginning of the two forms that triggers the divergent behaviour. I know of no such examples involving initial full or partial geminates. Nonetheless, this is arguably the situation in English with sC-clusters. Some instances of initial sC are preceded by zero; others are preceded by schwa.

(111) *Initial empty nuclei before sC-clusters in English*

a. With preceding zero [ØsC]	b. With preceding schwa [əsC]
state, steam, strange	*estate, esteem, estrange*
scarp, scape(goat)	*escarp, escape*
special, spy	*especial, espy*

It is not possible that magic licensing of initial empty nuclei is 'on' in English for the forms in (111a), but 'off' for those in (111b). However, these minimal pairs can be differentiated by a representational distinction. For example, the forms preceded by zero might start with an empty nucleus lacking both melody and a root node (112a), whereas the forms preceded by schwa might start with an empty nucleus including a root node but still no underlying melody (112b).

(112) a. *state* b. *estate*



The empty root node in (112b) triggers a default vowel. In English, this means automatic insertion of the @ element, which is realized as schwa.

Admittedly, the cross-linguistic evidence in favour of a distinction involving different degrees of degeneracy in initial empty nuclei is rather slim. This situation contrasts with the fairly abundant evidence uncovered regarding medial and even final empty nuclei. Also, although I have identified some apparent examples of initial empty nuclei with a default vowel, I have not found examples of initial empty nuclei whose content is determined by harmony. Under the circumstances, I have little choice but to restate what previous phonologists (e.g., Kaye 1992) have expressed: initial empty nuclei (whether before sC, NC or CC-clusters) are special. In languages that permit them, they show a preference for remaining empty. In the representational system I am proposing, this means that initial empty nuclei tend to lack a root node. I have no clear explanation for this distributional preference, although I tentatively speculate that the pattern serves partly to demarcate a crisp left edge.

Likewise, I cannot be sure why initial empty nuclei only occur with [s] or else with a partial or full geminate in coda position. That is, one does not find forms such as [Øktor] or [Øpteet], despite these initial consonant sequences being acceptable coda-onsets in some languages (e.g., in English *doctor* or *captive*). Essentially, there is something exceptional about sC, NC and CC, related to [s] and to partial and full geminates being generally preferred codas. The occurrence of partial and full geminates is arguably more understandable in this context, since so little (or none) of their content needs to be directly licensed by the coda position itself. When it comes to initial coda [s], however, like Kaye, I am obliged to appeal to prestidigitation.

4.5 Conclusion

The ability to employ degenerate forms such as empty nuclei with or without a root node is a universal option that only some languages select. In addition, degenerate forms can either be lexically specified or derived. For example, a filled nucleus may become an empty nucleus via a reduction process involving total suppression of element content, leaving only the timing slot or the timing slot plus root node. Processes of fortition and lenition may also apply to empty nuclei, either adding or removing a root node. When languages employ empty nuclei *with a root node* as part of the representational arsenal, these degenerate forms may be filled by default or by local spreading (harmony). In the case of empty nuclei *without a root node*, these may either remain empty or be filled by local spreading (copy epenthesis). Both types of local spreading thus constitute an option that a language is free to select or reject.

When harmony does not apply to an empty nucleus with a root node (e.g., in French, German and Palauan), insertion of a default vowel such as [ə] or [ɨ] is automatic. In essence, the *raison d'être* of the root node is to harbour elements, and the vacuum of an unspecified root node demands element content, triggering its insertion. When harmony does apply, element content is supplied by an adjacent position. For example, in Turkish, the elements I and U can spread from the previous nucleus. The fact that the target empty nucleus contains its own root node determines that it can harbour select individual elements from neighbouring positions.

Conversely, an empty nucleus without a root node lacks the requisite prosodic structure to anchor its own element content. It can neither generate automatic default content nor cherry-pick individual elements from its neighbours. Nonetheless, the presence of a timing slot leaves another option open. The role of a timing slot is to supply duration to the elements contained under a root node and to act as intermediary between the melodic tier and the tier of syllabic constituents. Consequently, an empty nucleus with a timing slot can potentially acquire an association line with a nearby root node. This is what occurs in Selayarese and Winnebago when the root node under an adjacent nucleus spreads to the available timing slot under the empty nucleus. The result is a doubly-linked root node, which generates a series of two identical vowels at the surface. Like individual element spreading (harmony), root node spreading (copy epenthesis) is not available in all languages. Turkish, for example, instantiates harmony involving individual elements in its empty nuclei with a root node; its empty nuclei without a root node are not subject to copy epenthesis, instead remaining empty.

The existence of two types of spreading provides particularly compelling evidence for the proposal that empty nuclei either contain or lack a root node. Given the presence of a root node, we would expect that, alongside default content, the empty nucleus should be capable of harbouring one or more individual elements supplied by neighbouring segments. Conversely, in the event that no root node is present, neither selective spreading of elements nor default insertion should be an option – the requisite prosodic unit for harbouring elements is no longer in place. The presence of a timing slot, on the other hand, opens the door for spreading of a neighbouring root node. This is precisely the dichotomy in behaviour we find with empty nuclei.

Interestingly, there appears to be a constraint on spreading to targets that lack a root node. Copy epenthesis applies only to a single subsequent nucleus, whereas vowel harmony can apply to successive nuclei (whether filled or empty). Apparently, a root node can be associated to timing slots in at most two nuclei, but individual elements can be linked to a whole series of root nodes.

GP attempts to account for the distribution of empty nuclei that are expressed or remain empty via the Empty Category Principle. In a nutshell, empty nuclei are required to be licensed to remain empty either by parameter or by proper government. We encountered, however, numerous cases in this chapter where empty nuclei could be expressed or fail to be realized in one and the same context. At times, empty nuclei are realized despite being licensed; at others, they remain empty despite not being licensed. Such data are entirely inconsistent with the Empty Category Principle, but they accord well with a representational approach based on differing degrees of degeneracy.

At times, whether a nucleus in a given form is filled or empty and, when empty, whether it contains a root node or not is simply an arbitrary lexical choice. This is largely the case in Turkish. Nonetheless, the distribution of filled and empty nuclei is not always unconstrained. In general, a correlation is found between degree of degeneracy and prosodic recessiveness. Frequently, stressed syllables are required to have full vowels, such that underlyingly empty nuclei may alternate with filled nuclei under stress (e.g., French *lever* [lə̲'ve] versus *lève* ['lɛ̲v]). Conversely, as we saw from vowel reduction in Palauan, either zero or default vowels such as schwa may alternate with full vowels under stress shift. Interestingly, there is a tendency for syllables with empty nuclei that lack a root node to be invisible for the purpose of stress placement (Selayarese) or syllable count (Trukese). Hypothetically, syllables with empty nuclei that lack a root node may not always be incorporated into foot structure (an issue which can be addressed in future research). Finally, empty nuclei are required to contain a root node (i.e., to be realized) in certain contexts for licensing purposes. Word-internally in French, for example, empty nuclei must contain a schwa after a branching onset or a coda-onset sequence. Apparently, only a nucleus with a root node confers on an onset a 'license to govern' a dependent (Charette 1990). This requirement is, however, relaxed word-finally, giving rise to alternations such as *libre̲ment* (with schwa) and *libr(e̲)* (with zero).

Word-initial empty nuclei before sC-clusters and full and partial geminates constitute a special case. Typically, these empty nuclei remain empty, suggesting that they lack a root node. I know of no cases of local spreading targeting such absolute initial nuclei, and cases where they are subject to default content are hard to find. In GP, these initial empty nuclei are particularly troubling because they remain empty under circumstances where empty nuclei are typically required to be realized (i.e., they are not properly governed). I cannot claim to have come up with a satisfactory explanation for these special cases, but I submit that they pose a thorny problem for any theoretical framework.

Now we turn to the behaviour of one last form of empty category, namely empty codas. These too exhibit behaviour consistent with an analysis based on two (or even three) degrees of degeneracy.

CHAPTER V

EMPTY CODAS

5.1 Introduction

As with empty onsets and empty nuclei, I will argue that empty codas (i.e., 'rhymal adjuncts') come in two varieties, corresponding to different degrees of degeneracy in the prosodic representation. Lexically, empty codas can contain either a timing slot plus root node devoid of element content or only a timing slot – see Figure 5.1. The former are typically realized as glottal stops; the latter as geminates. One difference between empty codas without a root node and their empty onset or nucleus counterparts is that empty codas are never empty at the surface level. Essentially, there would be no reason to posit an empty coda unless it is realized; given their enduring non-head status, codas have no licensing duties that would require their presence in a representation independently of some surface manifestation. Finally, coda nasals that are homorganic with an adjacent onset are in fact partially empty codas that are specified only for the nasal element N.

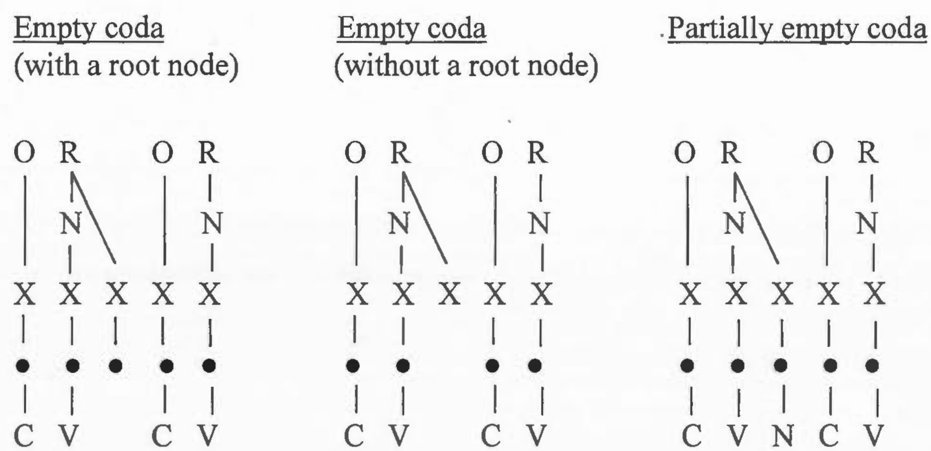


Figure 5.1 *Types of empty coda*

The latter two representations in Figure 5.1 should be fairly uncontroversial. An empty coda without a root node is the underlying representation corresponding to surface geminates such as the following from Italian and Japanese in (113).

(113) a. <u>Italian geminates</u> b. <u>Japanese geminates</u>

toppo	'stump'	*happa*	'leaf'
fatto	'fact'	*hatta*	'stuck'
bevve	'he drank'	*sakka*	'author'
palla	'ball'	*hasso*	'sending'
carro	'cart'	*suragga*	'slugger'
donna	'woman'	*kiddo*	'kid'

In these examples, all of the syllables closed by a consonant identical to the neighbouring onset are in fact represented lexically with an empty timing slot as rhymal dependent (Kaye, Lowenstamm & Vergnaud 1990). This empty timing slot is filled through regressive spreading of the root node in the adjacent onset. The spreading process creates a heterosyllabic long consonant, as shown in Figure 5.2.

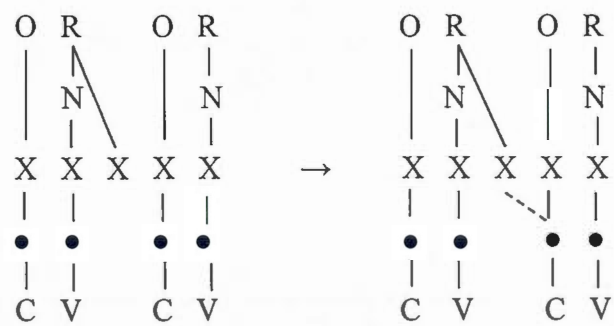


Figure 5.2 *Derivation of geminate consonant*

A partially empty coda is the representation for coda nasals that are homorganic to an adjacent onset. This construction is extremely common among languages that permit codas – indeed, so much so that it is hard to find coda languages that do not permit homorganic nasals in this location. Some examples are given below from Spanish, Japanese and Ojibwa.

(114) a. Spanish homorganic nasals

ca[m]*po*	'field'	*balu*[m]*bo*	'bulky thing'
ma[n]*to*	'cloak'	*ba*[n]*da*	'band'
ma[ŋ]*co*	'one-armed'	*ta*[ŋ]*go*	'tango'
a[m]*fora*	'amphora'	*ma*[n]*so*	'calm'

b. Japanese homorganic nasals

to[m]*bo*	'dragonfly'	*ka*[ŋ]*gae*	'thought'
si[n]*do-i*	'tired'	*u*[n]*zari*	'disgusted'

c. Ojibwa homorganic nasals

o[m]*bibide:*	'it flies up'	*ga:*[n]*ǰibin*	'push someone'
mi[n]*dido*	'he is big'	*ba*[ŋ]*gisin*	'it falls'

I assume the coda nasals in all of the above examples to be represented underlyingly as containing only the element N. The homorganicity of the nasals [m, n, ŋ] that appear in the surface forms in Spanish, Japanese and Ojibwa is derived via regressive spreading of place from the adjacent onset, a process illustrated in Figure 5.3.

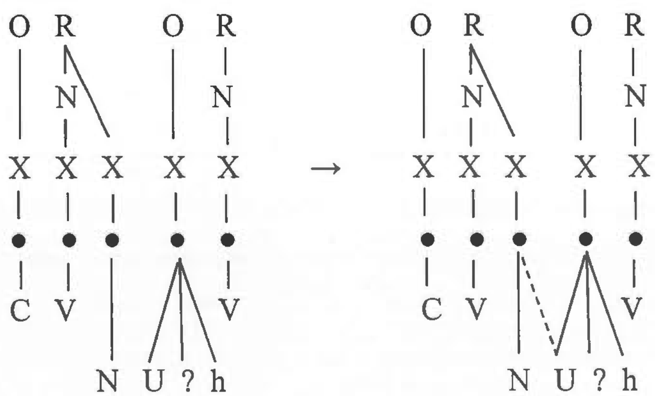


Figure 5.3 *Derivation of a homorganic coda nasal:* Nb → mb

In the above example, onset [b], composed of the elements U-?-h, supplies the coda nasal with a labial specification in the form of the element U. There is a problem with this portrait of the derivation of [mb] from [Nb], however: part of the melodic content of the nasal is unaccounted for. In the above derivation, the coda nasal starts with only the element N and ends with both N and U, but the segment [m] is composed of the elements N-U-?. Where does the missing occlusion element come from?

Two conceivable solutions might be: (i) to include the ?-element in the lexical specification of the coda nasal from the start, such that a coda nasal would already contain the elements N-?, rather than just N (a view expressed in Harris 1994); (ii) to derive the presence of ? in the coda nasal via an additional process of spreading from the onset – hence, in the case of Nb → mb, the onset stop would supply the coda nasal with both U and ?. To my mind, both of these solutions must be rejected. Instead, I propose a third solution: (iii) automatic insertion of the default element ? accounts for its presence in the coda nasal. But before presenting arguments in support of this view, I will outline the problems generated by the other two solutions.

An important part of the appeal of a representation of coda nasals as containing only the element N is that this way the coda has minimal melodic licensing responsibilities. Codas are notoriously poor melodic licensors (Ito 1986; Goldsmith 1990; Lombardi 1995b; Harris 1997), so if coda nasals only need to be specified for the element N, this goes a long way to explaining why nasals are universally preferred coda segments (Prince 1984). By treating coda nasals as lexically specified for both N and ? (the solution in (i) above), we undermine to a degree the justification for coda nasals. For this reason, such a solution to accounting for the presence of ? is undesirable.

By deriving the presence of ? via spreading from the onset stop, we avoid attributing greater licensing potential to the coda position (albeit at the cost of a more complex derivation). This is certainly a point in favour of the solution in (ii) above. On the other hand, we encounter problems when a homorganic nasal is followed by a fricative, as in Spanish *a*[m]*fora* or Japanese *u*[n]*zari*. The fricative onsets [f] and [z] contain the elements U-h and R-h respectively, but neither contains ?. Consequently, while the fricative can supply the coda nasal with the place specification U or R, it cannot provide the element for occlusion. Spreading from the adjacent onset thus cannot account for the presence of the element ? in the coda nasal.

Instead, I propose that occlusion is generated automatically in coda nasals via default insertion of the ?-element. A central aim of this chapter is to explore evidence in favour of this view. The revised derivation of coda nasals is illustrated in Figure 5.4.

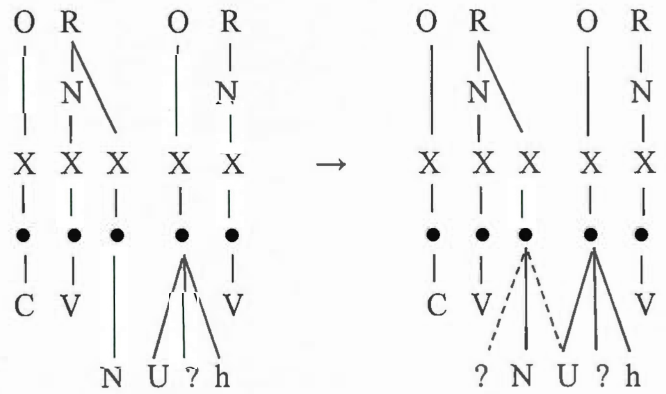


Figure 5.4 *Revised derivation of homorganic coda nasal:* Nb → mb

Another aim of this chapter is to argue in favour of two types of empty coda: those with and without a root node. As we have seen, those without a root node generate geminates: the root node of the adjacent onset associates via spreading to the empty timing slot of the coda. On the other hand, as with empty onsets, I propose that the presence of a root node in an empty coda typically triggers insertion of the ?-element. Consequently, a glottal stop appears automatically in an empty coda with unspecified root node as shown below in Figure 5.5. To my knowledge, glottal stops are universally quite rare in coda position, whereas geminates are comparatively common. In other words, for some reason, the more degenerate representation of an empty coda without a root node is preferred to the less degenerate form with a root node (unless this root node contains minimally the N-element). I have no explanation for why this should be so.

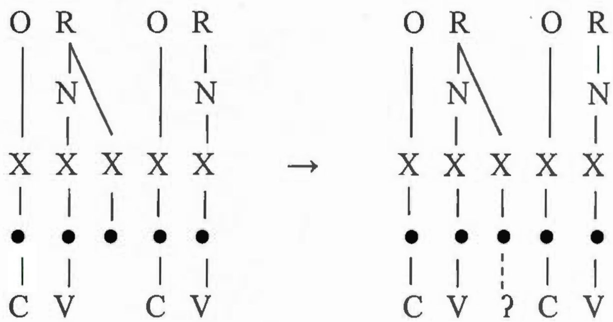


Figure 5.5 *Derivation of a coda glottal stop*

A final aim of this chapter is to argue, contra the orthodox GP position, that codas (whether empty or filled) can appear in three locations: word-initially after an empty nucleus and before an onset, word-internally before an onset, and word-finally – see Figure 5.6.

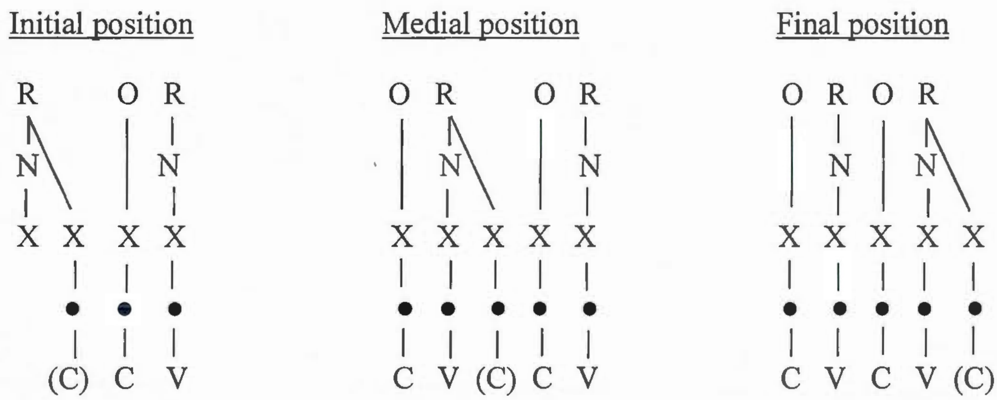


Figure 5.6 *Three locations for empty/filled codas*

The notion of final codas contravenes the Coda Licensing Principle proposed by Kaye (1990), which states that codas must always be licensed by an adjacent onset. Even assuming that final codas can at times be licensed across a word boundary by the onset of an following form fails to account for final codas that precede pauses.

Consequently, as outlined previously in Chapter I, GP assumes that final consonants are always onsets of empty nuclei, never codas. I will adopt the position defended by Piggott (1999, 2003b) and Rice (2003), however, that while in some languages final consonants are indeed onsets of empty nuclei, in others final consonants are codas.

Languages that instantiate final codas include Selayarese, Kiribatese, Lardil, Ojibwa, and Manam (according to Piggott), as well as Ahtna (according to Rice) and Brazilian Portuguese (in my view). The arguments adduced in support of a final coda analysis are essentially the inverse of the arguments presented earlier to support an onset analysis. First, in some languages, the same set of consonants appears in final position as in medial codas. Next, for purposes of stress assignment, final consonants sometimes contribute to the weight of the final syllable.

The next section presents an extended analysis one of the languages cited by Piggott as having final codas: Selayarese (documented in Mithun & Basri 1986, but see also Goldsmith 1990; McCarthy & Prince 1994; Basri 1999; Broselow 2000; Finer 2000; Basri, Broselow & Finer 2012). Building on Piggott's analysis, I demonstrate that Selayarese in fact employs two types of empty coda: those with and without a root node. In addition, Selayarese contains partially empty codas, lexically specified only for the element N. The occurrence of these three sorts of degenerate coda in Selayarese is partially delimited by the nature of the adjacent onset consonant, but their distribution is not otherwise constrained. That is, their selection is principally a matter of lexical choice, and no correlation is found between degree of degeneracy and relative prosodic or morphological prominence.

5.2 Empty codas in Selayarese

For reference, the inventory of underlying consonants for Selayarese is given below in Table 5.1.

Table 5.1 *Selayarese consonant inventory*

Stops	p b	t d	ɟ	k g	(ʔ)
Fricatives		s			h
Prenasalized stops	ᵐb	ⁿd	ᶮɟ	ᵑg	
Nasals	m	n	ɲ	ŋ	
Liquids		l r			

Although it is identified as an underlying consonant in Mithun and Basri (1986) and elsewhere, I have placed the glottal stop in parentheses because to my mind its status as lexically specified segment is debatable (on this, see also Goldsmith 1990). Prenasalized stops are found only in medial position, where they contrast with nasal-stop sequences: for example, [bo:ᵐbaŋ] 'wave' vs. [bombaŋ] 'bamboo skin used for binding'. They are also somewhat volatile across dialects: some speakers consistently replace them with a nasal-stop sequence or with a simple voiced stop. Otherwise, all consonants are found initially and medially in onsets, including glottal stops. Vowel-initial syllables also occur in both initial and medial position (e.g., [a:su] 'dog' and [ge:oʔ] 'deed, behaviour'). Interestingly, only glottal stop and the velar nasal [ŋ] appear word-finally.

The set of underlying vowels is limited to [i, e, a, o, u]. Final mid vowels are subject to surface laxing and trigger regressive laxing harmony (Basri & Chen 2000). Vowels are also subject to progressive nasal harmony after nasal consonants. Neither of these processes will be indicated in the data presented here. A process of vowel lengthening in stressed open syllables (the only context for long vowels) is indicated whenever pertinent (or simply whenever given in the source text). Stress is penultimate, and it shifts rightward as suffixes (though not suffixal clitics – Finer 2000) are added so as to maintain penultimate position. Consequently, we find [gó:lo] 'ball' but [goló:ku] 'my ball'. As was discussed in Chapter III, final epenthetic vowels in Selayarese are invisible for purposes of stress assignment. These vowels (determined harmonically by the preceding vowel) occur after roots that end underlyingly in [r], [l] or [s]. Hence, while [sahá:la] 'sea cucumber' with final specified vowel shows expected penultimate stress, [sá:hala] 'profit' with final epenthetic vowel exhibits antepenultimate stress.

I will argue here for two types of empty coda in Selayarese, corresponding to different degrees of degeneracy in the prosodic representation: empty codas may contain either a timing slot and root node or only a timing slot. In my view, the former are realized as glottal stops (whence the placement of glottal stop in parentheses above); the latter surface as geminates. Selayarese also has partially empty codas, lexically specified only for the N-element. These surface as nasals homorganic to an adjacent onset. No other form of coda is found in this language. I will also argue that codas appear in three locations in Selayarese: initially, medially and finally. Since languages often allow codas only in medial position, the varied distribution of codas in Selayarese (along with the range of empty coda types) makes this language a particularly rich object of study.

5.2.1 Geminates (empty codas without a root node)

Probably the least controversial empty codas in Selayarese are the ones responsible for geminate consonants: namely, empty codas without a root node. The distribution of this coda type is more limited than the other two: empty codas without a root node (and hence geminate consonants) are found in initial and medial position, but not in final position.

(115) *Empty codas without a root node (geminate consonants)*

a. Initial position

[ppallui]	'he/she is cooking'
[ttu:nui]	'he/she burned'

[kke:keki]	'you (hon) dig'
[ssinsiŋ]	'to wear a ring'
[hhattui]	'it is seasonal'
[mma:ri]	'to stop'
[nnoʔnoso]	'to shake a liquid'
[lla:ri]	'to run'

b. <u>Medial position</u>

[ruppa]	'face'
[ditte]	'me'
[kukkuru]	'to cut off all hair'
[sassa]	'to wash'
[ʔahhattui]	'it is seasonal'
[rammasa]	'rammasa'
[bannaŋ]	'thread'
[maŋŋaŋ]	'tired'
[paɲɲu]	'turtle'
[tallu]	'three'
[barroʔ]	'conceited'

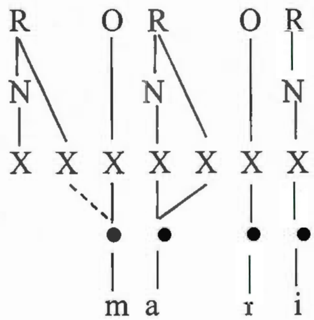
c. <u>Final position</u>

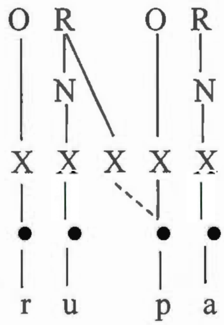
XX

The possibilities for gemination are more restricted in initial than in medial position. First, among liquids, geminate [rr] occurs only in medial position. Likewise, among nasals, geminate [ɲɲ] and [ŋŋ] are found only in medial position. Finally, most of the time only voiceless obstruents geminate in Selayarese. There are, however, a few exceptional forms involving the voiced stops [b] and [g] as medial geminates (e.g., [sabbara] 'to be patient' and [hugga] 'a kind of insect'), but never in initial position. Note also that geminate glottal stops *[ʔʔ] are not found in any position.

The differences between initial and medial positions in the distribution of geminates can be attributed to differences in the preceding nucleus. As argued in Chapter III, initial geminates are coda-onset sequences preceded by an empty nucleus (116a). Medial geminates of course are preceded by a filled nucleus (116b). Given that the nucleus serves as intraconstituent licensor to the coda position, it is not surprising that the less degenerate filled nucleus should support a greater range of geminate codas.

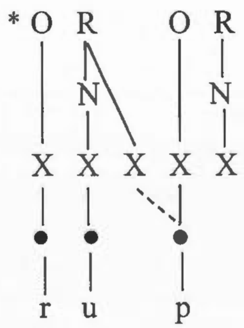
(116) a. <u>Initial geminate</u> b. <u>Medial geminate</u>

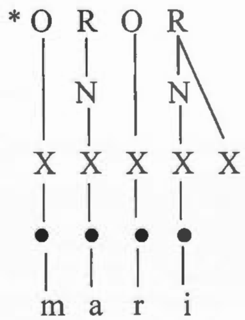




A third conceivable configuration, namely a final geminate (i.e., an identical coda-onset sequence followed by an empty nucleus without a root node), is unattested in Selayarese (117a). Likewise, there are no apparent instances of empty codas without a root node in absolute final position (i.e., without an adjacent onset) (117b).

(117) a. *<u>Final geminate</u> b. *<u>Final empty coda</u> (without root node)





The configuration in (117a) is not universally illicit.[83] Final geminates are found in some variants of Arabic, including the Hadhrami dialect of Yemen: [ˈrabb] 'Lord', [ʔaˈxaff] 'lighter' (Bamakramah 2009, cited in Davis 2011). Final geminates also occur in Wolof: [segg] 'to filter', [dəpp] 'to put upside down' (Bell 2003). In order for them to occur, a language must permit final empty nuclei that lack a root node and hence remain empty. This appears not to be the case for Selayarese. Only final empty nuclei with a root node occur, for example in forms such as [sá:hala] 'profit', where the final nucleus is filled via progressive spreading of the vowel from the previous nucleus: [sá:halØ] → [sá:hala].

The configuration in (117b) above is rather improbable, since there would be no reason to posit an empty coda in the absence of an adjacent onset that supplies content. Certainly, empty nuclei that remain empty may be posited if their presence is required for licensing purposes. Likewise, wherever there is a nucleus, an empty onset that remains empty may be posited since, being heads, onsets are obligatory constituents. Neither of these conditions applies to empty codas, however, which have no licensing responsibilities because they are enduring non-heads.

Nonetheless, the structure in (117b) is precisely what is involved in the process of *raddoppiamento sintattico* in Italian (discussed previously in Chapter III). Briefly, the phenomenon involves gemination of a word-initial onset following a word that ends in a short stressed vowel. Thus, *caffé fréddo* 'iced coffee' is realized as [kafféf**f**réddo]. The process serves to make the stressed syllable heavy by supplying it with a coda. Arguably, the final empty coda can be portrayed as either derived whenever a consonant-initial word follows the stressed short vowel or as underlyingly present.

[83] Why this configuration is illicit in Selayarese but not other languages is an open question. Hypothetically, in Selayarese, final empty nuclei without a root node do not license the preceding onset to govern a coda dependent; in other languages, the empty nucleus does confer the requisite licensing potential. This situation closely parallels the distinction between French and English referred to in note 59, Chapter IV: unexpressed final empty nuclei can license a branching onset in French (e.g., [tablØ] 'table') but not in English (e.g., [tejbəl]/*[tejblØ]).

Determining which is not crucial for our purposes, but the absence of compensatory lengthening of the short vowel whenever a consonant-initial word does not follow probably favours the derivational view.

5.2.2 Partial geminates (partially empty codas)

Partially empty codas definitely occur in initial and medial position, as attested by the homorganic nasal + obstruent/liquid sequences in (118a,b) below. In final position, we do not find any NC-sequences, though we do find the velar nasal [ŋ] in absolute final position (118c) to the exclusion of any other consonants except [ʔ]. Clearly, to treat final [ŋ] as an instance of a partially empty coda (as I do here) is more contentious, so this view will need to be justified.

(118) *Codas with only the N-element (partial geminates)*

a. Initial position

[mbannaŋ]	'to stretch'
[nda:tala]	'to pursue'
[nta:le]	'to cross'
[nsu:lu]	'to get out'
[nri:o]	'to hold'
[ɲɟa:ma]	'to work'
[ŋgo:ra]	'to shout, cheer'
[ŋha:u]	'to kiss'

b. Medial position

[sambaŋ]	'to trip'
[lumpa]	'to jump'
[ʔondaŋ]	'to invite'
[nontoŋ]	'to watch'

[ʔansu:luʔ]	'to get out'
[kanre]	'food'
[ɟaɲɟaŋ]	'to see'
[taŋgoŋ]	'to guarantee'
[miŋka]	'but'
[ʔaŋhukkaŋ]	'to punish'

c. Absolute final position

[ŋa:seŋ]	'all'
[ŋi:nuŋ]	'drink'
[maŋŋaŋ]	'tired'

As with geminates, a more restricted set of partial geminates is permitted in initial versus medial sites. The restriction involves sequences of a nasal + voiceless stop: in medial position, [mp], [nt] and [ŋk] are found, whereas in initial position we find only [nt]. Inexplicably, [nl] fails to appear in any position. Interestingly, note that a glottal stop never appears after a nasal. Arguably, this might be due to the fact that glottal stops (composed of just the ?-element) lack a place element. Note however that velar [ŋ] appears both initially and medially before [h], even though [h] (composed of just the h-element) also lacks place. Finally, the only nasal to appear in absolute final position is likewise velar [ŋ].

I analyze this final [ŋ] as having the same representation as the coda of initial and medial partial geminates. Namely, it is a partially empty coda, specified only for the element N. In its surface form as [ŋ], of course, it contains the elements N-?-@. Clearly, there is a need to justify the analysis of final [ŋ] as: i) occupying a coda, a view that violates the Coda Licensing Principle (Kaye 1990) since there is no adjacent onset to license the coda; and ii) being partially empty (i.e., lexically specified only for N, rather than for N-?-@).

One phenomenon that might point to final nasals occupying a coda is their pervasive assimilation to an adjacent consonant either within a reduplicated form (119a) or across a word boundary (119b).

(119) *Final nasal assimilation*

a. In reduplication:

Root		Reduplicated form	
[pekaŋ]	'hook'	[pekampekaŋ]	'hook-like object'
[tunruŋ]	'hit'	[tunruntunruŋ]	'hit lightly'
[ɟaŋaŋ]	'chicken'	[ɟaŋaɲɟaŋaŋ]	'bird'
[gintaŋ]	'chili'	[gintaŋgintaŋ]	'chili-like object'
[soroŋ]	'push'	[soronsoroŋ]	'drawer'
[hukkuŋ]	'punish'	[hukkuŋhukkuŋ]	'punish lightly'
[roŋgaŋ]	'loose'	[roŋganroŋgaŋ]	'rather loose'
[lamuŋ]	'grow'	[lamullamuŋ]	'plantation'

b. Across word boundaries:

[annaŋ]	'six'	[annaŋ hoʔoŋ]	'six erasers'
[annam baɟu]	'six shirts'	[annan rupa]	'six kinds'
[annan dareʔ]	'six monkeys'	[annal loka]	'six bananas'
[annaɲ ɟaraŋ]	'six horses'	[annam mata]	'six eyes'
[annaŋ koko]	'six gardens'	[annan nolo]	'six zeroes'
[annan sapo]	'six houses'	[annaɲ ɲaha]	'six souls'

In assimilating to the place of the adjacent consonant, final nasals behave exactly like word-initial or internal nasals in partial geminates. The most straightforward conclusion is that final nasals are of the same ilk: they are coda consonants.[84] As the

[84] This conclusion is not unassailable. Final nasals in English (which I analyze as onsets of empty nuclei), for example, also assimilate across word boundaries: *i*[n] *Tokyo*, *i*[m] *Paris*, *i*[ŋ] *Cambridge*, *i*[ɱ] *fact*, *i*[ɲ] *church* (Spencer 1996). However, these assimilations are variable and, arguably, they operate at the level of phonetic implementation, where syllable structure may no longer be relevant.

reduplicated form [lamullamuŋ] shows, when [ŋ] and [l] come together, the outcome of contact is geminate [ll]. Again, this follows if the nasal is in a coda: in initial and medial partial geminates, nasals do not appear before [l] (i.e., *[nl]), but geminate [ll] is found in initial and medial positions. If the final [ŋ] were simply the onset of an empty nucleus, there is no reason the constraint on nasal/lateral-liquid adjacency should apply. Likewise, gemination results when [ŋ] is adjacent to another nasal across the word boundary in the forms [annam mata], [annan nolo] and [annaɲ ɲaha]. This behaviour is also consistent with a coda analysis.

The coda analysis is significantly bolstered by the behaviour of nasal-final forms before consonant-initial suffixes (or 'affixal clitics' – Basri 1999) such as the personal markers *-ku* 'my', *-mu* 'your', *-ba* 'our' (excl), *-ta* 'our' (incl), *-na* 'his/her/its'. In this context, as expected, the nasal assimilates to the adjacent suffix consonant. In addition, as a result of suffixation, stress shifts to the (now penultimate) stem-final syllable, but crucially the vowel fails to lengthen (120).[85]

(120) *Nasal assimilation and absence of vowel lengthening in stem-final syllables*

gó:tiŋ	-ba	→	[gotímba]	'our (excl) scissors'
gó:tiŋ	-ta	→	[gotínta]	'our (incl) scissors'
ʔá:suŋ	-ku	→	[ʔasúŋku]	'my dog'
mántaŋ	-na	→	[mantánna]	'his/her/its eye'

Nasal assimilation could conceivably take place even if the NC-sequence is parsed as a pair of onsets with an intervening empty nucleus (see note 84). The absence of vowel lengthening, on the other hand, constitutes behaviour consistent only with a coda parse for the nasal. That is, if the nasal occupies a coda, these are precisely the syllabic conditions that obviate vowel-lengthening.

[85] Examples in the literature are not abundant, but these few here are indicative of the pattern.

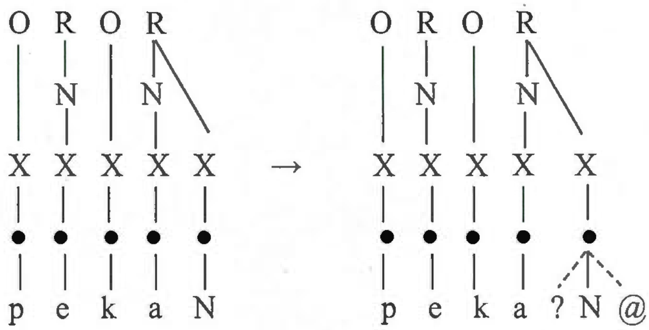
Next, I address the question of whether final nasals are lexically specified only for the N-element. First, nasals in onsets clearly must be fully specified. For example, there is no local source for the place elements in the initial nasal onsets in the forms [nikka] 'marriage', [mo:noʔ] 'stop', [ɲa:ha] 'soul', and [ŋa:seŋ] 'all'. Also, given the range of place elements that appear in the forms, place cannot be attributed to a default setting. Consequently, place must be part of the lexical representation of onsets. In final position, on the other hand, only velar nasals occur. Place is not contrastive in this context, and there is no local source for velarity. We may thus posit that velar place is assigned by default: the element @ is automatically inserted into a nasal coda with unspecified place.[86] The fact that velarity is represented by the element @, precisely the element that appears by default in empty nuclei, makes the hypothesis of default insertion in nasal codas all the more plausible. In contrast, if the element @ is deemed to be specified in the lexical representation of final nasals, this would impose an inordinate licensing burden on a position that lacks the requisite strength.

It is pertinent that Selayarese also exhibits [-ATR] harmony: final mid vowels lax and laxing then spreads to preceding mid vowels (for more on this, see Basri & Chen 2000). That is, from an element perspective, the element @ is inserted into a final nucleus and spreads regressively, in a manner similar to Quebec French laxing harmony (Poliquin 2006). For my purposes, this laxing behaviour is significant because it shows that the element @ is independently required to have default epenthetic status in Selayarese nuclei. The extension of default @-insertion to nasal codas is thus particularly plausible.

Based on the previous arguments, I propose that final velar nasals are codas specified only for the element N. The further two elements @ (for velar place) and ? (for occlusion) that, in the absence of place assimilation, comprise the surface form are derived via default insertion into the partially specified root node.

[86] For a related view on the unmarkedness of velar place, see Rice (1996).

(121) *Derivation of surface final* [ŋ]: [pekaN] → [pekaŋ]



A final argument in favour of the view I am promoting is based on the occurrence of [ŋ] word-initially and internally in coda position before [h]: [ŋha:u], [ʔaŋhukkaŋ]. The glottal fricative contains only the noise element h. Consequently, this segment is incapable of determining the place specification of an adjacent nasal. One possibility, then, would be that exceptionally, in this context before [h], [ŋ] is lexically specified for all three of its surface elements: N-@-?. In this case, however, a violation of the Complexity Condition as it applies to coda-onset sequences occurs: onset [h] with its single element is insufficiently complex to government-license coda [ŋ] with its three elements. If the nasal is only specified for N, on the other hand, no such violation occurs. In my view, then, the elements ? and @ must both be assigned to initial and medial coda [ŋ] by default before [h]. Default insertion of these elements is thus not limited to final position in Selayarese.

In sum, partial geminates technically occur only in initial position (following an empty nucleus) and in medial position (following a filled nucleus). It has been argued here that the initial member of the partial geminate is a coda specified only for the N-element. The additional place element is acquired from the adjacent onset, and the occlusion element ? from a process of default insertion. The final velar nasal [ŋ], however, has the same representation as the initial member of an NC-sequence: it is a coda specified only for N. As with the nasal coda in true partial geminates, final nasals acquire the occlusion element ? by default; in addition, they are assigned unmarked velar place in the form of the default element @. The same default place emerges in initial and medial coda nasals before [h]. The analysis of final nasals as codas is controversial in that it contravenes the GP Coda Licensing Principle. Arguably, what allows the final nasal to circumvent Coda Licensing is its ability to acquire place by default. That is, in languages where Coda Licensing holds and final consonants (including nasals) are all onsets, final nasals resist default insertion of place, absolutely requiring it to be supplied by an adjacent onset.

5.2.3 Glottal stops (empty codas with a root node)

The only other consonant to appear in first position in initial and medial consonant sequences is [ʔ]. Glottal stops appear in this position before a wide range of consonants (122a,b). They also appear in absolute final position (122c). Following Piggott (2003b), I adopt the view that these glottal stops are parsed as codas. I further propose that these are not lexically specified segments; rather they are the surface realization of an empty coda with a root node.

(122) *Empty codas with a root node (glottal stop)*

a. Initial position

[ʔba:lasa]	'to reciprocate' (intr) [87]
[ʔdeŋka]	'to pound with a mortar'
[ʔɟa:ri]	'to become' (intr)
[ʔganraŋ]	'to play the drum'
[ʔma:ru]	'to share a husband'
[ʔnikka]	'to marry' (intr)
[ʔŋo:aʔ]	'to yawn' (intr)[88]

[87] A number of the initial glottal codas in these examples are derived from the intransitive prefix [aʔ], which surfaces as [ʔ] due to initial vowel loss preceding a stressed syllable (Mithun & Basri 1986; Basri 1999). These forms are thus morphologically complex. To my knowledge, however, a morpheme boundary is not always present between an initial glottal stop and the following consonant.

[ʔlampa]	'to go' (intr)
[ʔra:ga]	'to play rattan ball'

b. <u>Medial position</u>

[laʔba]	'lack of salt'
[ʔaʔdappeʔ]	'to fall'
[paʔɟa]	'salty'
[taʔga:raŋ]	'be stained'
[kaʔmu:ru]	'nose'
[ɟeʔneʔ]	'water'
[t:aʔɲuʔɲuʔ]	'to stumble'
[ʔaʔŋo:aʔ]	'to yawn' (intr)
[seʔla]	'salt'
[seʔre]	'one'

c. <u>Absolute final position</u>

[se:keʔ]	'tight'
[sassaʔ]	'lizard'
[to:doʔ]	'also'

Glottal stops never appear before voiceless obstruents. This is the inverse of what we found with geminates: normally, only voiceless obstruents geminate (with a few exceptional instances of medial [bb] and [gg]). In the case of glottal stops, the ban on appearing beside voiceless obstruents is exceptionless. Indeed, the ban is so complete that prefixes ending in glottal stop trigger gemination when attached to roots beginning with a voiceless obstruent: see for example the allomorphs of the stative verb intransitivizer prefix [taʔ-] (123).

---

[88] The gloss 'to yawn' (intr) is given for both [ʔŋo:aʔ] (17a) and [ʔaʔŋo:aʔ] (17b) in Mithun and Basri. Though these are the only examples of initial and medial [ʔŋ] sequences provided, the authors explicitly state that [ʔŋ] occur both initially and medially so presumably further examples can be found in the language.

(123) *Gemination across morpheme boundaries:* [taʔ-]

a. Before vowel-initial roots		b. Before sonorant-initial roots	
[taʔataʔ]	'to be roofed'	[taʔmuri]	'to smile'
[taʔenteŋ]	'to be erected'	[taʔlesaŋ]	'to be removed'
[taʔinuŋ]	'to be drunk (liquid)'	[taʔrinriŋ]	'to be walled'

c. Before voiceless obstruents		d. Before voiced obstruents	
[tappelaʔ]	'to get lost'	[taʔbessolo]	'to slip'
[tattuda]	'to bump against'	[taʔdoʔdoʔ]	'to be sleepy'
[takkaluppa]	'to faint'	[taʔɟai]	'to be sewn'
[tassambaŋ]	'to stumble/trip'	[taʔgaraŋ]	'to get strained'

Gemination also takes place across word boundaries within an intonation group when a word ending in a glottal stop (e.g., [appaʔ] 'four') abuts a word beginning with a voiceless obstruent (124a); however, the glottal stop is able to surface before a voiced obstruent or sonorant (124b).

(124) *Gemination across word boundaries:* [appaʔ] 'four'

a. Before voiceless obstruents

[appaʔ#pao]	→	[appappao]	'four mangoes'
[appaʔ#tajuʔ]	→	[appattajuʔ]	'four flowers'
[appaʔ#kura]	→	[appakkura]	'four small turtles'
[appaʔ#sapo]	→	[appassapo]	'four houses'
[appaʔ#hugga]	→	[appahhugga]	'four insects'

b. Before voiced obstruents/sonorants

[appaʔ#balo]	→	[appaʔbalo]	'four holes'
[appaʔ#dareʔ]	→	[appaʔdareʔ]	'four monkeys'
[appaʔ#golo]	→	[appaʔgolo]	'four balls'

[appaʔ#mata]	→	[appaʔmata]	'four eyes'
[appaʔ#loka]	→	[appaʔloka]	'four bananas'

Much the same questions that were posed regarding final [ŋ] can be applied to the view advanced here that glottal stops in Selayarese may appear in empty codas with a root node:

1. Why are pre-consonantal and final glottal stops treated as occupying a coda? Could they not in fact be onsets? Hypothetically, they might form either the first part of a branching onset or the onset of an empty nucleus. The analysis as onset of empty nucleus holds particular appeal for glottal stops in final position since parsing them in a coda incurs a violation of the Coda Licensing Principle (Kaye 1990).
2. Why are glottal stops treated as empty categories in Selayarese? Could they not be underlying segments that are lexically specified for the ʔ-element?
3. If pre-consonantal and final glottal stops are in fact empty codas, what distinguishes them from the empty codas responsible for geminates? Could they not have the same representation?

Regarding question (1), there is no reason to suspect that ʔC-sequences might be branching onsets. First, Selayarese does not have typical branching onsets (i.e., stop + liquid sequences such as [bl] or [dr]). Next, alongside ʔC, any sequences that do occur are canonical coda-onsets (i.e., as we have seen, geminates such as [kk] or [ss], and partial geminates such as [mb] or [nr]). Finally, according to the Complexity Condition, a segment such as [ʔ] with a single element ʔ cannot be the head a branching onset.

Before we treat ʔC as incontrovertibly another instantiation of a coda-onset pair, however, we might consider whether the glottal stop could not be an onset of an unexpressed nucleus (i.e., an empty nucleus without a root node as were posited in

French, German, Palauan, and Turkish in Chapter III). Initial and medial ʔC in, say, [ʔdeŋka] and [laʔba] could in fact be a sequence of two onsets separated by an empty nucleus: [ʔØ.deŋ.ka] and [la.ʔØ.ba]. Likewise final glottal stops as in [se:keʔ] might be followed by an empty nucleus that allows the glottal stop to be parsed into an onset: [se:.ke.ʔØ]. Certainly, under this analysis, we avoid violating the Coda Licensing Principle.

There are reasons to reject this latter proposal. First, as shown previously, the vowels in stressed (penultimate) syllables lengthen in open syllables but not in closed syllables. Hence [tó:doʔ] and [llá:ri], but [r**úp**pa] and [l**úm**pa] – the stressed vowel lengthens except when the syllable is closed by a geminate or partial geminate coda. Likewise, when the stressed vowel is followed by a ʔC-sequence, lengthening fails to apply: [**láʔ**ba], [s**éʔ**la]. The inevitable conclusion is that ʔC-sequences have the same prosodic representation as geminates and partial geminates: namely, they are coda-onsets (but cf. Kaye 1990 for a view that this logic is not unassailable).

Relatedly, when a consonant-initial suffix is added to a form ending in a glottal stop such that the final syllable is now in stressed penultimate position, the vowel fails to lengthen (125a). Final glottal stops thus exhibit the same behaviour as the final nasals shown before in (120). This holds whether the glottal stop is preserved in the complex form, as in [sassáʔmu], or whether it assimilates to a voiceless obstruent [sassákku]. When the same suffixes are added to a vowel-final root, on the other hand, the stressed vowel does lengthen (125b).

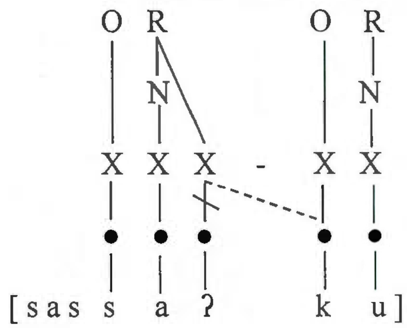
(125) *Final glottal stop/final vowel with consonant-initial suffixes*

a. Final glottal stop		b. Final vowel	
[sássaʔ]	'lizard'	[bá:ɟu]	'shirt'
[sassákku]	'my lizard'	[baɟú:ku]	'my shirt'
[sassáʔmu]	'your (fam) lizard'	[baɟú:mu]	'your (fam) shirt'

[sassátta]	'our (incl) lizard'	[baɟú:ta]	'our (incl) shirt'
[sassáʔba]	'our (excl) lizard'	[baɟú:ba]	'our (excl) shirt'
[sassáʔna]	'his/her/its lizard'	[baɟú:na]	'his/her/its shirt'

It is revealing that, as with the earlier examples in (123) and (124), the glottal stop does not surface beside a voiceless obstruent; instead, a geminate consonant is created. This alternation only provides further support for the view that the final glottal stop itself is in a coda, since coda position is precisely the context for gemination. Certainly, if the glottal stop is analyzed as an empty coda with a root node, all that is needed to generate the context for gemination is to delink the root node (126).

(126) *Generation of geminate* CC *from* ʔC: [sássaʔ-ku] → [sassákku]



In brief, it makes more sense to analyze pre-consonantal and final glottal stops as codas. This is not to say, of course, that final glottal stops always remain in coda position. Presumably, the glottal stop in prefix [taʔ-] re-associates to the empty onset of a vowel-initial root as in [taʔataʔ] – see (123) above. Similarly, addition of the absolutive clitic [-i] to [ʔ]-final roots such as [lumpaʔ] 'to jump' and [bakkaʔ] 'big' (127) likely involves re-association of the glottal stop to the empty onset.[89]

[89] One complication that I do not address here is the difference in behaviour of final [ʔ] before vowel-initial true suffixes as opposed to suffixal clitics as in (127). Before the transitive true suffix [-i] or the comparative true suffix [-aŋ], not only does the glottal stop re-associate to onset position, but it also alternates with [k]: [lumpaʔ-i] → [lumpaki] 'to jump across'; [bakkaʔ-aŋ] → [bakkakaŋ] 'bigger'. On the one hand, this might lead us to surmise that, in absolute final position at least, glottal stops are

(127) *Re-association of final coda* [ʔ] *to adjacent empty onset*

Before absolutive clitic [-i]

[ʔlumpaʔi]	'he jumped'
[bakkaʔi]	'it is big'

Nonetheless, it seems fully justified to otherwise parse pre-consonantal and final glottal stop into coda position. In addition, I assume that preconsonantal and final glottal stops are empty codas with an unspecified root node. This leads in to the second question that needs to be addressed: Why are glottal stops treated as empty categories in Selayarese? Could they not be underlying segments that are lexically specified for the ?-element?

A number of facts about glottal stops in Selayarese suggest that they are empty categories rather than lexically filled segments and, more precisely, that they are empty categories with an unspecified root node. This holds whether the glottal stop is associated to an onset or to a coda constituent. First, as we saw in sections 5.2.2 and 5.2.3, an onset glottal stop cannot license a coda: glottal stops do not geminate (i.e., *[ʔ.ʔ]), nor are they in second position of a partial geminate (i.e., *[N.ʔ]). Of course, glottal stops are simplex segments, composed only of the ?-element, so we may be inclined to attribute their inability to license a coda to their lack of complexity. This line of reasoning proves unjustified, however, when we consider the behaviour of the glottal fricative [h]. Despite also being simplex, composed only of the noise element h, [h] is capable of licensing a coda, whether in a geminate [h.h] or in a partial geminate [N.h]. This behaviour is doubly revealing since typically stops are better onsets than fricatives. Indeed, some languages lack fricatives (e.g., Lardil – Hale

underlying [k]. Certainly, Mithun and Basri state that historically there was an automatic alternation between preconsonantal [ʔ] and prevocalic [k], such that [ʔ] was at one time an allophone of [k]. This is no longer the case since prefix [taʔ-] maintains its glottal stop before vowel-initial roots such as [taʔ-enteŋ] 'to be erected'. Likewise, [ʔ] and [k] are clearly contrastive in minimal pairs such as [saʔa] 'snake' and [saka] 'hoof'. On the other hand, then, we might treat forms with [k] as suppletive, constituting traces of a historical situation where surface glottal stop was indeed an allophone of [k]. More research is required to be certain, but I suspect this to be the correct conclusion.

1973), but to my knowledge none lacks stops. The mismatch in licensing ability between [h] and [ʔ] makes sense, however, if we posit a fundamental difference in their representation: [h] is a lexically specified segment, whereas [ʔ] is the realization of an empty category.

Further support for the view of glottal stops in Selayarese as empty categories rather than as fully specified segments comes from consonant epenthesis. In a number of contexts, glottal stops are used as epenthetic segments (whether in onset or in coda position). Reasonably, their preferred status in epenthesis is due to their minimal structure: in order to generate a glottal stop in an onset or coda, all that is required is a timing slot with an unspecified root node, which is automatically supplied with the ʔ-element. The contexts for glottal epenthesis include:

(i) between identical vowels;
(ii) before vowel-initial forms at the beginning of an intonation group;
(iii) in certain reduplicated forms.

Identical vowels cannot be separated by an empty onset that lacks a root node in Selayarese. Within monomorphemic roots, we never find adjacent identical vowels; instead, identical vowels are always separated either by a bona fide lexically specified consonant (a filled onset) or by a glottal stop (an empty onset with a root node). When identical vowels come into contact due to morphological concatenation, a glottal stop is inserted. For example, when the first person absolutive marker [-a] is attached to a root ending in [a] or when the third person absolutive marker [-i] is attached to a root ending in [i], a glottal stop appears; when the root ends in a non-identical vowel, however, no glottal stop is present (128).

(128) *Glottal stop epenthesis between identical vowels*

[halli]	'buy'	[liŋka]	'walk'
[amall**ia**]	'I bought'	[aʔliŋk**aʔa**]	'I walked'

[amall**iʔi**]	'(s)he bought'	[aʔliŋk**ai**]	'(s)he walked'

The same pattern is found when identical vowel sequences arise through attachment of a vowel-final prefix to a vowel-initial word (Mithun & Basri 1986). Likewise, when identical medial vowel sequences are generated through loanword adaptation, a glottal stop is inserted (Broselow 2000).

Glottal stops also appear in empty onsets before vowel-initial forms at the beginning of an intonation group. For example, the word [inni] 'this' is pronounced with an initial glottal stop [ʔinni] after a pause, but without the glottal stop within an intonation group: [ʔa:pa inni] 'what is this?'. The same utterance-initial fortition process is of course found in other languages, including English and Japanese (Vance 1987). In Chapter III, also, particularly with respect to German, it was argued that glottal stop epenthesis in onsets is best analyzed simply as default insertion of the element ʔ into an unspecified root node. The same appears to apply in Selayarese, except that such unspecified root nodes can appear in both onset and coda position.

Glottal stop epenthesis in coda position also occurs in certain reduplicative forms. Selayarese reduplication targets the first two syllables of a root (and there are no monosyllabic roots). In the simplest cases, bisyllabic roots that end in a vowel, [ŋ] or [ʔ] are reproduced as exact copies in the reduplicant, which appears in initial position (129a,b,c). As shown previously in (119), [ŋ] at the end of a reduplicant assimilates in place to the adjacent onset or else it becomes a geminate before [l]. Likewise [ʔ] before a voiceless obstruent becomes a geminate.[90]

[90] Incidentally, one implication of the data in (129) is that, if reduplication truly does target only the first two syllables of a root, final velar nasals and glottal stops are part of the second syllable and hence clearly codas. The behaviour thus reinforces the coda parse advanced so far.

(129) *Reduplication of bisyllabic roots*

a. <u>Vowel-final roots</u>

[bahi]	'pig'	[bahibahi]	'toy pig'
[golla]	'sugar'	[gollagolla]	'candy'

b. <u>Nasal-final roots</u>

[pekaŋ]	'hook'	[pekampekaŋ]	'hook-like object'
[tunruŋ]	'hit'	[tunruntunruŋ]	'hit lightly'
[lamuŋ]	'grow'	[lamullamuŋ]	'plantation'

c. <u>Glottal-stop-final roots</u>

[toboʔ]	'to stab'	[tobottoboʔ]	'to stab repeatedly'
[pelaʔ]	'to throw away'	[pelappelaʔ]	'to throw away repeatedly'

When polysyllabic roots are reduplicated, only the first two syllables are copied. Interestingly, the second syllable of the reduplicant is always closed by a glottal stop, which assimilates in typical fashion to an adjacent voiceless onset.

(130) *Reduplication of polysyllabic roots*

[balala]	'greedy'	[balaʔbalala]	'rather greedy'
[palola]	'eggplant'	[paloppalola]	'eggplant-like object'
[kalammaŋ]	'dark'	[kalakkalammaŋ]	'rather dark'
[kaloko]	'porridge'	[kalokkaloko]	'porridge-like object'

The motivation for glottal stop epenthesis in this context is not entirely clear. Nonetheless, the point to retain is that it is a glottal stop that can be inserted by default when required. Coupled with their inability to geminate *[ʔ.ʔ] or to support a partial geminate nasal coda *[N.ʔ], the cumulative impression from the various processes of glottal stop epenthesis, whether in an onset or a coda, is that glottal stops in Selayarese are not simply lexically specified segments like any other.[91] They have special default status which sets them apart. In a nutshell, their behaviour is consistent with an analysis of glottal stops as generated from unspecified root nodes.

The third question that was raised concerns whether a distinction is really required between the empty coda that generates a default glottal stop and the empty coda that generates a geminate. Could a single degenerate representation not underlie both forms of surface coda? Certainly, before obstruent onsets, glottal stop and geminate codas appear in mutually exclusive contexts: glottal stops occur only with voiced obstruents, and geminates (almost) only with voiceless obstruents.[92] Given this distribution, the occurrence of the two types of degenerate coda is predictable, and a representational distinction is not necessary. The problem arises before sonorant onsets, since both glottal stops and geminates can appear in this context: [ka**ʔm**u:ru] versus [ra**mm**asa]. Given that glottal stops and geminates are contrastive in this context, a representational distinction must be in place. The distinction I have argued for involves two degrees of degeneracy: either the empty coda contains a timing slot and unspecified root node, in which case content in the form of the default element ʔ is automatically supplied, or else it contains only a timing slot, in which case content is supplied by regressive spreading of the root node under the adjacent onset. The former configuration generates a glottal stop, and the latter a geminate consonant.

---

[91] Another context in which glottal stops can appear out of nowhere in a coda is following an epenthetic vowel in a form such as [sáhala] 'profit' before a consonant-initial suffix such as [-na] 'his/her/its': [sahaláʔna].

[92] The reason for this complementary distribution is not immediately apparent. Piggott (2003b) attributes the situation to a constraint (Voice Licensing) that excludes voiced obstruent geminates: "A voiced obstruent in non-head position is unlicensed." Nonetheless, this constraint does not capture the whole picture. First, exceptions such as [sabbara] and [hugga] show that voiced obstruents are not strictly barred from codas in Selayarese. Second, as we have amply observed, geminate codas alternate with glottal stops before voiceless obstruents. This alternation suggests that it is more the context glottal stop + voiceless obstruent that needs to be constrained. Unfortunately, any constraint I could propose would be (like Piggott's own constraint) merely stipulative.

5.3 Conclusion

To reiterate, geminates [CC] and glottal stop codas [ʔC] can appear in the same context in Selayarese, namely with a sonorant in second position. Their overlapping distribution necessitates a representational distinction to contrast the two degenerate forms. This chapter proposes that geminates and glottal stop codas constitute two types of empty coda. Like empty nuclei and empty onsets, empty codas may have two degrees of degradation in their representation. First, they may be simply timing slots associated to a rhyme, in which case they trigger gemination: the root node under the adjacent onset spreads regressively, creating a doubly linked structure and hence a long consonant. In Selayarese, this form of empty coda appears in initial position following an empty nucleus or in medial position following a filled nucleus, but not in final position. In final position, by definition there is no adjacent onset to supply a degenerate coda with content, so true geminate codas are barred from this context. Second, empty codas may have both a timing slot and a root node, a representation which triggers insertion of the default element ? and thus glottal stop production. Essentially, the unspecified root node, the role of which is to group together melodic elements, automatically sanctions the presence of the ?-element in a coda. In Selayarese, this form of empty coda appears in initial, medial and final positions.

Furthermore, codas may be partially empty, minimally specified for just the nasal element N. This representation leads to production of a homorganic nasal in initial and medial positions and to a default velar nasal [ŋ] in final position. In all three positions, the occlusion element ? is generated by default in the nasal coda. In initial and medial positions, place is determined by spreading from the adjacent consonant. The exception is before [h], in which case velar place is assigned by default insertion of the element @. In final position, velar place is also determined by default. The insertion of @ without cost in a coda mirrors its behaviour in empty nuclei with a root node, where the element @ triggers schwa production.

The representation of final [ʔ] and [ŋ] as codas challenges the strictness of the Coda Licensing Principle. Undeniably, codas are in general followed and hence licensed by an onset, so the claim is not that the principle has no validity. However, I have shown that Selayarese permits empty codas with a root node to surface in final position without licensing from an adjacent onset. The root node in the empty coda may be completely unspecified, leading to a glottal stop, or specified only for N, leading to a velar nasal. In sum, the findings support the view expressed in Piggott (1999, 2003b) and Rice (2003) that not all languages adhere to Coda Licensing.

CONCLUSION

When apparently identical phonological forms such as empty onsets, nuclei and codas behave differently in identical contexts, this divergent behaviour poses a serious challenge to grammatical and functional approaches alike. A language cannot employ one parameter setting or constraint ranking some of the time, but another the rest of the time. Likewise, if one attributes the behaviour of a phonological form to articulatory ease or perceptual salience, this behaviour should not arbitrarily change from one lexical item to another. Essentially, then, behavioural differences in apparently identical forms constitute the problem to which the representational distinction I propose is a solution. Specifically, the representation of empty syllabic categories (onsets, nuclei and codas) can differ based on the presence versus absence of an unspecified root node. This abstract distinction, which can be underlying or derived, has crucial repercussions for how the empty category behaves.

Within GP, a representational distinction has in fact been adopted to account for the differing behaviour of *h-aspiré* and vowel-initial words in French: the former supposedly begin with an empty onset containing a timing slot, whereas the latter start with an empty onset that lacks a timing slot. In Chapter III, it was shown that this particular means of configuring the representational distinction encounters considerable problems. Specifically, it fails to predict the differences in behaviour that are actually found. For example, the distinction cannot motivate the application of liaison to vowel-initial forms versus the blocking of liaison with *h-aspiré* forms. Instead, then, I propose to distinguish between the two types of empty onset based on the presence versus absence of a root node. This particular distinction proves to be an effective predictor of the behaviour that is found.

It is not only empty onsets, however, that exhibit differing behaviour in identical contexts: empty nuclei and codas show the same variation. In general, differences in the behaviour of empty categories can be divided into: i) whether a default segment can be inserted or not; and ii) whether individual elements or only entire segments can spread from adjacent positions. To my knowledge, this distinction between two forms of spreading has not been identified previously in the literature, much less accounted for.

According to my proposal, when an empty category contains a root node, it can either harbour default content or be the target of selective spreading of elements. Essentially this behaviour falls out from the role of a root node in representations, so the choice of the root node to make the distinction is by no means arbitrary. The root node serves as a docking site, grouping together the different melodic elements and linking them to the timing tier. When a root node is unspecified, it either generates elements such as @ and ? by default (realized as schwa and glottal stop respectively) or else it attracts one or more elements from neighbouring segments. Although the exact quality of default content can vary from one language to another, its insertion appears to be automatic, while filling of the unspecified root node via spreading is an option that only some languages select.

When empty onsets with a root node are the target of spreading, typically the glides [j] or [w] are formed by spreading of the elements I or U from an adjacent segment. Because the target is endowed with its own root node, the adjacent source segment may be either simplex (such as [i] and [u]) or complex (such as [e] and [o]). This follows because the glide is formed via double association of the element itself to two root nodes. When empty onsets without a root node are the target of spreading, however, glide formation is necessarily limited to simplex sources. Without a target root node, glide formation entails double association of the source root node with the target timing slot. Furthermore, in the absence of spreading, the empty onset without a root node simply remains empty.

The same distinction applies to empty nuclei. Those with a root node can generate default content such as schwa, whereas those without a root node cannot. Individual elements can form double associations with a source and target root node (i.e., vowel harmony). Otherwise, when the target empty nucleus lacks a root node, spreading must implicate the root node under a neighbouring timing slot. In this case, the spreading operation will always share elements from a neighbouring position in their entirety (i.e., copy epenthesis).

Empty codas are a bit of a special case. First, their presence is not required in a representation for licensing purposes, so they are only posited when they surface as filled. Next, evidence for the presence of empty codas with a root node is relatively rare. Selayarese employs glottal stop by default in coda position, a phenomenon that can be attributed to the presence of an empty coda with a root node. On the other hand, I have not been able to uncover instances where an empty coda with a root node is the target of spreading. Typically, spreading takes the form of gemination. In this case, the target is an empty coda without a root node: the root node under an adjacent onset spreads to the timing slot under the empty coda, forming a long consonant. Additionally, spreading can target partially empty codas, that is, codas that are lexically specified only for the N element. Such underspecified codas receive their place specification via spreading from an adjacent onset, and the occlusion element ? appears by default. The output of spreading is thus a homorganic nasal stop. When no adjacent onset is available, in Selayarese at least, the velar place element @ is assigned by default. Indeed, throughout the preceding chapters, a recurring property of the elements @ and ? has been an ability to appear in a representation from nowhere. Importantly, the view defended here that empty codas (with partially or entirely default content) can appear in final position challenges the strictness of the

Coda Licensing Principle, whereby a coda under normal circumstances must be licensed by an adjacent onset.

In a nutshell, part of the representational arsenal that is universally available (albeit not always utilized in a language) involves constructing empty categories of two degrees of degeneracy: either with or without a root node. In certain cases, the choice to use either type of empty category, or indeed a filled category, is purely lexical. This is not always the case, however.

In GP, it has been proposed that the distribution of realized and unrealized empty nuclei is determined by the Empty Category Principle. This principle states that an empty nucleus must be licensed in order to remain empty. Depending on the context, licensing of an empty nucleus is determined either by parameter or by proper government from an adjacent realized nucleus. In numerous examples from various languages, however, it was demonstrated that the distribution of empty nuclei simply does not consistently conform to the patterns predicted by the proposed licensing mechanism. At times, empty nuclei must be realized despite being licensed to remain empty; at others, empty nuclei remain empty despite not being licensed. Under the proposal based on a representational distinction, however, the observed patterns are unproblematic: the realization of an empty nucleus is crucially linked to the presence versus absence of a root node, not to whether an empty nucleus is licensed to remain empty. Nonetheless, this is not to say that the distribution of empty nuclei with or without a root node is entirely unconstrained.

A general correlation is found between the degree of degeneracy of a representation and prosodic or morphological prominence. For example, there is a tendency for stressed syllables and/or syllables at the left edge of content words to require either a filled onset or an empty onset that includes a root node. Empty onsets without a root node may be relegated to unstressed syllables and syllables at the beginning of function words. The same general pattern emerges with respect to nuclei. Filled nuclei are frequently required in stressed syllables; empty nuclei, particularly those that lack a root node, tend to be limited to unstressed syllables. As a result of this constraint on their distribution, languages may show alternations between filled and empty nuclei as a result of stress shift. Similarly, languages may require that a stressed syllable contain an empty coda (with or without a root node). When this empty coda is realized, whether by default or via spreading, it confers requisite weight to the stressed syllable. With all three types of empty category, then, a consistent pattern is for prominent positions to require less degenerate forms.

That not all languages employ empty categories suggests that their presence is somehow determined by parameter (or else by constraint ranking, from an OT perspective). I have already proposed an empty coda parameter ("Codas may be empty"), which accounts for the fact that not all languages with codas permit geminates, despite Prince's (1984) assertion that these are among the least marked codas. Similar parameters can surely be proposed for empty onsets and nuclei. In future research, I plan to explore this issue in greater detail. For the moment, suffice it to say that it seems likely that separate parameters are required to govern empty categories with and without unspecified root nodes. Indeed, it may prove fruitful to propose a markedness relation, with empty onsets and nuclei without a root node being more marked than those with a root node. This would explain notably why, in English second language acquisition, learners frequently realize final empty nuclei first with an epenthetic default vowel before being able to maintain them as null (Broselow et al. 1998; Broselow 2004; Broselow & Xu 2004). In other words, this behaviour is indicative of a developmental pathway where empty nuclei with a root node are acquired (or posited) before empty nuclei without a root node. Similarly, since there is a general preference for empty nuclei in domain-final sites, this option may also be determined by parameter, a proposal that echoes the GP view that final nuclei are licensed to remain empty via parameter.

On a final note, more work (inevitably) needs to be done on certain issues that have been addressed only partially or simply left unanswered. First, I consider that my account of the distribution of the two types of empty category and their filled counterparts is still too general. Demonstrating that the mechanisms proposed by GP are ultimately inadequate is only a first step. What is needed now is a principled account that explains precisely why a particular representation is required in one context or, conversely, barred from another. Next, I have not explicitly addressed the question of why certain elements spread to empty categories in a given system, while others do not. When an empty onset with a root node is the target, it makes sense that the elements I and U spread from an adjacent nucleus, since these generate the unmarked glides [j] and [w]. Spreading of the elements A or @ would not have the same effect. Why Turkish vowel harmony should also be restricted to the spreading of I and U, on the other hand, is not immediately apparent. Likewise, why harmony that targets A-nuclei should only implicate the I-element constitutes another mystery. Whether such patterns can be captured by other means than simple stipulation remains then an open question. For these reasons, really this conclusion marks another beginning.

REFERENCES

Alber, B. (2001). Regional variation at edges: Glottal stop epenthesis in standard and southern varieties of German. *Zeitschrift für Sprachwissenschaft, 20,* 3-41.

Alderete, J. (1995).Winnebago accent and Dorsey`s Law. In J.N. Beckman, L.W. Dickey, & S. Urbanczyk (Eds.), *Papers in Optimality Theory, University of Massachusetts Occasional Papers 18, (pp. 21-51).* Amherst, MA: GLSA.

Anderson, J.M., & Ewen, C.J. (Eds.). (1980). *Studies in dependency phonology. (Ludwigsburg Studies in Language and Linguistics, 4).* Ludwigsburg: R.U.O. Strauch.

Anderson, J.M., & & Ewen, C.J. (1987). *Principles of dependency phonology.* Cambridge: Cambridge University Press.

Anderson, J.M., & Jones, C. (1974). Three theses concerning phonological representations. *Journal of Linguistics, 10,* 1-26.

Anderson, S.R. (1982). The analysis of French shwa: Or, how to get something from nothing. *Language, 58*(3), 534-573.

Antilla, A. (1997). Deriving variation from grammar: A study of Finnish genitives. In F. Hinskens, R. Van Hout, & L. Wetzels (Eds.), *Variation, change and phonological theory* (pp.35-68). Amsterdam: John Benjamins.

Archangeli, D., & Pulleyblank, D. (1994). *Grounded phonology.* Cambridge, MA: MIT Press.

Backley, P. (2011). *An introduction to Element Theory.* Edinburgh: Edinburgh University Press.

Bamakramah, M. (2009). *Syllable structure in Arabic varieties with a focus on superheavy syllables.* PhD dissertation, Indiana University.

Banksira, D.P. (2000). *Sound mutations: The morphophonology of Chaha.* Amsterdam/ Philadelphia: John Benjamins.

Barnes, J. (2006). *Strength and weakness at the interface: Positional neutralization in phonetics and phonology.* Berlin & New York: Mouton de Gruyter.

Basbøll H. (1974). Structure consonantique du mot italien. *Revue romane, IX,* 27-40.

Basri, H. (1999). *Phonological and syntactic reflections of the morphological structure of Selayarese.* PhD dissertation, State University of New York at Stony Brook.

Basri, H., Broselow, E., & Finer, D. (2012). The end of the word in Makassar languages. In T. Borowsky, S. Kawahara, T. Shinya & M. Sugahara

(Eds.), *Prosody matters: Essays in honor of Lisa Selkirk* (pp. 111-147). London: Equinox Publishing Ltd.

Basri, H., & Chen, Y. (2000). When harmony meets reduplication in Selayarese. In C. Kitto & C. Smallwood (Eds.), *AFLA VI: Proceedings of the Sixth Austronesian Formal Linguistics Association* (pp. 37-49). University of Toronto.

Bell, A.J. (2003). Gemination, degemination and moraic structure in Wolof. In *Working Papers of the Cornell Phonetics laboratory, 15,* 1-68.

Blaho, S. (2008). *The syntax of phonology: A radically substance-free approach.* PhD dissertation, University of Tromsø.

Boersma, P. (1998). *Functional Phonology: Formalizing the interactions between articulatory and perceptual drives.* The Hague: Holland Academic Graphics.

Boersma, P. (2005). Some listener-oriented accounts of hache aspire in French. Ms, University of Amsterdam, ROA #730.

Boersma, P., & Hayes, B. (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry, 32*(1), 45-86.

Booij, G.E. (1995). *The phonology of Dutch.* Oxford: Oxford University Press.

Brame, M.K. (1972). On the abstractness of phonology: Maltese ʕ. In M.K. Brame (Ed.), *Contributions to generative phonology* (pp. 22-61). Austin: University of Texas Press.

Brandon, F. (1975). A constraint on deletion in Swahili. In *The Proceedings of the Sixth Conference on African Linguistics. Ohio State Working Papers in Linguistics, 20,* 241-259.

Brockhaus, W. (1995). *Final devoicing in the phonology of German.* Tübingen: Niemeyer.

Broselow, E. (2000). Stress, epenthesis, and segment transformation in Selayarese loans. In S. Chang, L. Liaw, & J. Ruppenhofer (Eds.), *Proceedings of the 25th Annual Meeting of the Berkeley Linguistics Society* (pp. 311-325). Berkeley: University of California.

Broselow, E. (2004). Unmarked structures and emergent rankings in second language phonology. *International Journal of Bilingualism, 8* (1), 51-65.

Broselow, E., Chen, S.-I., & Wang, C. (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition, 20,* 261-280.

Broselow, E., & Xu, Z. (2004). Differential difficulty in the acquisition of second language phonology. *International Journal of English Studies, 4* (2), 135-163.

Cardoso, W. (2009). When input frequency fails to drive learning: The acquisition of sC onset clusters. In B. Baptista, A. Rauber, & M. Watkins (Eds.), *Recent research in second language phonetics/phonology: Perception and production* (pp. 67-85). Newcastle Upon Tyne: Cambridge Scholars Publishing.

Carlisle, R. (1991). The influence of environment in vowel epenthesis in Spanish/English interphonology. *Applied Linguistics, 12*(1), 76-95.

Casali, R.F. (1997). Vowel elision in hiatus contexts: Which vowel goes? *Language, 73*(3), 493-533.

Cedergren, H.J., & Sankoff, D. (1974). Variable rules: Performance as a statistical reflection of competence. *Language, 50*(2), 333-355.

Charette, M. (1985). The appendix in parametric phonology. In R. Schuh (Ed.)., *Studies in African Linguistics,* supplement 9, UCLA.

Charette, M. (1990). Licence to govern. *Phonology, 7*, 233-253.

Charette, M. (1991). *Conditions on phonological government.* Cambridge: Cambridge University Press.

Charette, M. (2006). The end of the (Turkish) word. *SOAS Working Papers in Linguistics, 14,* 23-40.

Charette, M., & Göksel, A. (1996). Licensing constraints and vowel harmony in Turkic languages. *SOAS Working Papers in Linguistics and Phonetics, 6*, 1-25.

Chierchia, G. (1986). Length, syllabification and the phonological cycle in Italian. *Journal of Italian Linguistics, 8,* 5-34.

Chomsky, N. (1981). *Lectures on government and binding: The Pisa lectures.* Berlin/New York: Mouton de Gruyter.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English.* New York: Harper & Row.

Clements, G.N. (1985a). Akan vowel harmony: A nonlinear analysis. In D.L. Goyvaerts (Ed.), *African linguistics. Essays in memory of M/W.K. Semikenke* (pp. 55-98). Amsterdam/Philadelphia: John Benjamins.

Clements, G.N. (1985b). The geometry of phonological features. *Phonology Yearbook, 2,* 225-252.

Clements, G.N. (1986). Compensatory lengthening and consonant gemination in LuGanda. In L.Wetzels & E. Sezer (Eds.), *Studies in compensatory lengthening* (pp. 35-77). Dordrecht: Foris.

Clements, G.N., & Hume, E.V. (1995). The internal organization of speech sounds. In J. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 245-306). Oxford: Blackwell.

Clements, G.N., & Keyser, S.J. (1983). *CV phonology: A generative theory of the syllable.* Cambridge, MA: MIT Press.

Clements, G.N., & Sezer, E. (1982). Vowel and consonant disharmony in Turkish. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations, Part 2* (pp. 213-257). Dordrecht: Foris.

Comrie, B. (1997). Turkish phonology. In A.S. Kaye (Ed), *Phonologies of Asia and Africa (part II)* (pp. 883–898). Winona Lake, IN: Eisenbrauns.

Côté, M.-H. (2007). Empty elements in schwa, liaison and h-*aspiré*: The French Holy Trinity revisited. In J.M. Hartmann, V. Hegedus & H. van Riemsdijk (Eds.), *Sounds of silence: Empty elements in syntax and phonology* (pp. 61-103). Oxford: Elsevier.

Crosswhite, K.M. (2000). Vowel reduction in Russian: A unified account of standard, dialectal, and 'dissimilative' patterns. University of Rochester Working Papers in the Language Sciences, 1(1), 107–172.

Cyran, E. (1996). Licensing properties of nuclei and principle ranking in Irish. *The Linguistic Review, 13,* 1-31.

Davis, S. (1984). Squamish stress clash. *Coyote Papers: Working Papers in Linguistics, 5,* 2-18.

Davis, S. (2011). Geminates. In M. van Oostendorp, C.J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology, Volume 2* (pp. 837-859). Malden, MA & Oxford: Wiley-Blackwell.

Davis, S., & Baertsch, K. (2012). Formal versus functional explanation for a universal theory of syllable structure: The case of vowel epenthesis in Winnebago. *Journal of Universal Language, 13*(2), 7-34.

Davis, S., & Cho, M.-H. (2003). The distribution of aspirated stops and /h/ in American English and Korean: An alignment approach with typological implications. *Linguistics, 41*(4), 607-652.

Dell, F. (1973). *Les règles et les sons: Introduction à la phonologie générative.* Paris: Hermann.

Downing, L.J. (2005). On the ambiguous segmental status of nasals in homorganic NC sequences. In M. van Oostendorp & J.M. van de Weijer (Eds.), *The internal organization of phonological segments* (pp. 183-209). Berlin: Mouton de Gruyter.

Durand, J. (Ed.). (1986a). *Dependency and non-linear phonology.* London: Croom Helm.

Durand, J. (1986b). French liaison, floating segments and other matters in a dependency framework. In J. Durand (Ed.), *Dependency and non-linear phonology* (pp. 161-201). London: Croom Helm.

Durand, J. (1987). On the phonological status of glides: The evidence from Malay. In J. Anderson & J. Durand (Eds.), *Explorations in Dependency Phonology* (pp. 79-107). Dordrecht: Foris.

Durand, J. (1990). *Generative and non-linear phonology.* London/New York: Longman.

Durand, J. (1995). Universalism in phonology: Atoms, structures and derivations. In: J. Durand & F. Katamba (Eds.), *Frontiers of phonology: Atoms, structures, derivations* (pp. 267-88). Longman, London.

Durand, J., & Eychenne, J. (2007). Remarks on schwa deletion and epenthesis in French. In *Actes des 5èmes Journées d'Études Linguistiques,* 89-94. Downloaded 29/07/13 from http://www.lling.fr/actes/actes-jel2007.pdf.

Durand, J., Slater, S., & Wise, H. (1987). Observations on schwa in southern French. *Linguistics, 25,* 983-1004.

Encrevé, P. (1988). *La liaison avec et sans enchainement.* Paris: Éditions du Seuil.

Ewen, C.J. (1995). Dependency relations in phonology. In J.A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 570-85). Oxford: Basil Blackwell.

Eychenne, J. (2009). Une variété de français conservatrice en Languedoc. In J. Durand, B. Laks, & C. Lyche (Eds.), *Phonologie, variation et accents du français* (pp. 259-284). Paris: Lavoisier.

Eychenne, J., & Pustka, E. (2007). The initial position in Southern French: Elision, suppletion, emergence. In *Actes des 5èmes Journées d'Études Linguistiques,* 199-204. Downloaded 29/07/13 from http://www.lling.fr/actes/actes-jel2007.pdf.

Finer, D. (2000). Cyclic clitics in Selayarese. In C. Kitto & C. Smallwood (Eds.), *AFLA VI: Proceedings of the Sixth Austronesian Formal Linguistics Association* (pp. 137-160). University of Toronto.

Flora, J.-A. (1974). *Palauan phonology and morphology.* PhD. dissertation, University of California, San Diego.

Foley, W.A. (1986). *The Papuan languages of New Guinea.* Cambridge: Cambridge University Press.

Fox, A. (2005). *The structure of German.* Oxford: Oxford University Press.

Fudge, E.C. (1969). Syllables. *Journal of Linguistics, 3,* 253–286.

Fudge, E.C. (1976). Phonotactics and the syllable. In A. Juilland (Ed.), *Linguistic studies offered to Joseph Greenberg* (pp. 381–398). Saratoga, CA: Anma Libri.

Gabriel, C., & Meisenburg, T. (2009). Silent onsets? An optimality-theoretic approach to *h aspiré* words. In C. Féry, F. Kügler & R. van de Vijver (Eds.), *Variation and gradience in phonetics and phonology* (pp. 163-184). Berlin: Mouton de Gruyter.

Gendron, J.-D. (1984). *Phonétique orthophonique à l'usage des Canadiens français.* Québec: Presses de l'université Laval.

Giegerich, H.J. (1985). *Metrical phonology and phonological structure: German and English.* Cambridge: Cambridge University Press.

Goad, H. (2012). sC Clusters are (almost always) coda-initial. *The Linguistic Review, 29*(3), 335-373.

Göksel, A., & Kerslake, C. (2005). *Turkish: A comprehensive grammar.* London/New York: Routledge.

Goldsmith, J.A. (1976/1979). *Autosegmental phonology.* PhD Dissertation, MIT. New York: Garland Press.

Goldsmith, J.A. (1990). *Autosegmental and metrical phonology.* Oxford: Basil Blackwell.

Guerssel, M. (To appear). On the licensing of glides. In S. Bendjaballah, M. Lahrouchi, N. Faust & N. Lampitelli (Eds.), *The form of structure, the structure of form: Essays on the realization of linguistic structures.* Amsterdam: John Benjamins.

Gussmann, E., & Kaye, J. (1993). Polish notes from a Dubrovnik café. I: The yers. *SOAS Working Papers in Linguistics and Phonetics, 3,* 427-462.

Hale, K. (1973). Deep-surface canonical disparities in relation to analysis and change: An Australian example. *Current Trends in Linguistics, 11,* 401-458.

Hale, K., & White Eagle, J. (1980). A preliminary metrical account of Winnebago accent. *International Journal of American Linguistics, 46,* 117-132.

Hale, M., & Reiss, C. (2000). Phonology as cognition. In N. Burton-Roberts, P. Carr, & G. Docherty (Eds.), *Phonological knowledge: Conceptual and empirical issues* (pp. 161-184). Oxford: Oxford University Press.

Hall, C. (1992). *Modern German pronunciation.* Manchester: Manchester University Press.

Hall, N. (2011). Vowel epenthesis. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (Eds.), *The Blackwell companion to phonology* (pp. 1576-1596). Malden, MA & Oxford: Wiley-Blackwell.

Hall, T.A. (1989). German syllabification, the velar nasal, and the representation of schwa. *Linguistics, 27,* 807-842.

Halle, M., & Vergnaud, J.-R. (1980). Three-dimensional phonology. *Journal of Linguistic Research, 1,* 83-105.

Halle, M., & Vergnaud, J.-R. (1987). *An essay on stress.* Cambridge, MA: MIT Press. Harris, J. (1997). Licensing Inheritance: An integrated theory of neutralisation. *Phonology, 14,* 315-370.

Harris, J. (1990). Segmental complexity and phonological government. *Phonology, 7,* 255-300.

Harris, J. (1994). *English sound structure.* Oxford: Blackwell.

Harris, J. (1997). Licensing Inheritance: An integrated theory of neutralisation. *Phonology, 14,* 315-370.

Harris, J., & Gussmann, E. (1998). Final codas: Why the west was wrong. In E. Cyran (Ed.), *Structure and interpretation in phonology: Studies in phonology* (pp. 139-162). Lublin: Folia.

Harris, J., & Gussmann, E. (2002). Word-final onsets. *UCL Working Papers in Linguistics, 14,* 1-42.

Harris, J., & Lindsey, G. (1995). The elements of phonological representation. In Jacques Durand & Francis Katamba (Eds.). *Frontiers of phonology: Atoms, structures, derivations* (pp. 34-79). Harlow, Essex: Longman.

Harris, J., & Lindsey, G. (2000). Vowel patterns in mind and sound. In N. Burton-Roberts, P. Carr, G. Docherty (Eds.), *Phonological knowledge: Conceptual and empirical issues* (pp.185-205). Oxford: Oxford University Press.

Harrison, D.K., & Kaun, A. (2001). Patterns, pervasive patterns and feature specification. In T.A. Hall (Ed.), *Distinctive feature theory* (pp. 211-236). Berlin: Mouton de Gruyter.

Hart, M.E. (1991). The moraic status of initial geminates in Trukese. In *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on the Grammar of Event Structure,* 107-120.

Hayes, B. (1981). *A metrical theory of stress rules.* PhD Dissertation, MIT, Cambridge.

Hayes, B. (1982). Extrametricality and English stress. *Linguistic Inquiry, 13,* 227-276.

Hayes, B., Kirchner, R., & Steriade, D. (Eds.). (2004). *Phonetically based phonology.* Cambridge: Cambridge University Press.

Herbert, R.K. (1975). Reanalyzing prenasalized consonants. *Studies in African Linguistics, 6* (2), 105-124.

Honeybone, P. (1999). 'I blame the government'. *Language Sciences, 21,* 177-221.

Hooper, J.B. (1972). The syllable in phonological theory. *Language, 48*(3), 525-40.

Hooper, J.B. (1976). *An introduction to natural generative phonology.* New York: Academic Press.

Hualde, J.I. (1992). Empty consonants and other CV-tier effects in Aranese Gascon. *Lingua, 86,* 177-188.

Hulst, H. van der. (1989). Atoms of segmental structure: Components, gestures and dependency. *Phonology, 6,* 253-284.

Hulst, H. van der. (2003). Cognitive phonology. In *Germania at alia. A linguistic webschrift for Hans den Besten on the occasion of his 55th birthday.*

Hulst, H. van der, and Ritter, N.A. (1999). Head-driven phonology. In Harry van der Hulst & Nancy A. Ritter (Eds.), *The syllable: Views and facts* (pp. 113–167). Berlin/New York: Mouton de Gruyter.

Hulst, H. van der, & Weijer, J. van de. (1991). Topics in Turkish phonology. In H. Boeschoten & L. Verhoeven (Eds). *Turkish linguistics today* (pp. 11-59). Leiden: E.J. Brill.

Ingleby, M., & Brockhaus, W. (2002). Phonological primes: Cues and acoustic signatures. In J. Durand & B. Laks (eds.), *Phonology: From phonetics to cognition* (pp. 131-150). Oxford: Oxford University Press.

Inkelas, S. (1999). Exceptional stress-attracting suffixes in Turkish: Representations versus the grammar. In R. Kager, H. van der Hulst & W. Zonneveld (Eds.), *The prosody–morphology interface* (pp. 134–187). Cambridge: Cambridge University Press.

Ito, J. (1986). *Syllable theory in prosodic phonology.* PhD dissertation, University of Massachusetts. Distributed by Indiana University Linguistics Club. Published 1988, New York: Garland.

Itō, J., & Mester, R.A. (1995). Japanese phonology. In J.A. Goldsmith (Ed.), The Handbook of Phonological Theory (pp. 817–838). Oxford: Blackwell Publishers.

Jones, D. (1956). English pronouncing dictionary. London: J.M. Dent & Sons.

Josephs, L. (1975). Palauan reference grammar. Honolulu: University of Hawaii Press.

Josephs, L. (1990). New Palauan-English dictionary. Honolulu: University of Hawaii Press.

Kabak, B., & I. Vogel (2001). The phonological word and stress assignment in Turkish. *Phonology, 18,* 315–360.

Kahn, D. (1976). *Syllable-based generalizations in English phonology.* PhD dissertation, MIT. Distributed by Indiana University Linguistics Club. Published 1980, New York: Garland.

Kassin, T.A. (2000). *The phonological word in standard Malay.* PhD dissertation, University of Newcastle.

Kawahara, S. (2003). On a certain kind of hiatus resolution in Japanese. *Onin Kenkyuu, 6*, 11-20.

Kawahara, S. (2004). Locality in echo epenthesis: Comparison with reduplication. In *Proceedings of NELS 34*, 295-309.

Kaye, J. (1985). On the syllable structure of certain West African languages. In Didier Goyvaerts (Ed.), *African linguistics: Essays in memory of M.W.K. Semikenke* (pp. 285-308). Amsterdam: John Benjamins.

Kaye, J. (1989). *Phonology: A cognitive view*. Hillsdale: Erlbaum.

Kaye, J. (1990). 'Coda' licensing. *Phonology, 7*, 301-330.

Kaye, J. (1992). Do you believe in magic? The story of *s+C* sequences. *SOAS Working Papers in Linguistics and Phonetics, 2*, 293-313.

Kaye, J. (2000). A user's guide to Government Phonology. MS, University of Ulster.

Kaye, J., & Lowenstamm, J. (1984). De la syllabicité. In F. Dell, J. Hirst & J.-R. Vergnaud (Eds.), *Forme sonore du langage* (pp. 123-160). Paris: Herman.

Kaye, J., Lowenstamm, J., & Vergnaud, R. (1985). The internal structure of phonological elements: A theory of charm and government. *Phonology Yearbook, 2*, 305-328.

Kaye, J., Lowenstamm, J., & Vergnaud, R. (1990). Constituent structure and government in phonology. *Phonology, 7*, 193-231.

Kiddo, M.E. (1993). *Tangale phonology: A descriptive analysis.* Berlin: Dietrich Reimer Verlag.

Kingston, J., & Beckman, M. (Eds.). (1990). *Papers in Laboratory Phonology I: Between the grammar and the physics of speech.* Cambridge: Cambridge University Press.

Kiparsky, P. (1968). How abstract is phonology? Bloomington: Indiana University Linguistics Club.

Kiparsky, P. (1993). Variable rules. *Paper presented at the Rutgers Optimality Workshop*. New Brunswick, NJ.

Kitto, C., & de Lacy, P. (1999). Correspondence and epenthetic quality. In C. Kitto & C. Smallwood (Eds.), *Proceedings of AFLA VI* (pp 181-200). Holland: Holland Academic Graphics.

Kohler, K.J. (1994). Glottal stops and glottalization in German. *Phonetica, 51,* 38-51.

Kornfilt, J. (1986). Stem-penultimate empty Cs, compensatory lengthening, and vowel epenthesis in Turkish. In L. Wetzels & E. Sezer (Eds), *Studies in compensatory lengthening* (pp. 79-96). Dordrecht: Foris.

Krämer, M. (2003). *Vowel harmony and Correspondence Theory.* Berlin: Mouton de Gruyter.

Lees, R. (1961). *The phonology of modern standard Turkish.* The Hague: Mouton & Co.

Levin, J. (1985). *A metrical theory of syllabicity.* PhD dissertation, MIT.

Lewis, G.L. (1967). *Turkish grammar.* Oxford: Oxford University Press.

Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry, 8,* 249-336.

Lombardi, L. (1995a). Laryngeal features and privativity. *The Linguistic Review, 12*, 35-59.

Lombardi, L. (1995b). Laryngeal neutralization and syllable wellformedness. *Natural Language and Linguistic Theory, 13*(1), 39-74.

Lombardi, L. (1996). Postlexical rules and the status of privative features. *Phonology, 13,* 1-38.

Lombardi, L. (2002a). Coronal epenthesis and markedness. *Phonology, 19*, 219-251.

Lombardi, L. (2002b). Markedness and the typology of epenthetic segments. In *Proceedings of Linguistics and Phonetics 2002.*

Major, R. (1986). Paragoge and degree of foreign accent in Brazilian English. *Second Language Research, 2,* 53-71.

Maris, M.Y. (1980). *The Malay sound system.* Kuala Lumpur Penerbit Fajar Bakti Sdn. Bhd.

Martlett, S.A., & Stemberger, J.P. (1983). Empty consonants in Seri. *Linguistic Inquiry, 14*(4), 617-639.

McCarthy, J.J. (1979/1985). *Formal problems in Semitic phonology and morphology.* PhD Dissertation, MIT. New York: Garland Press.

McCarthy, J.J. (1988). Feature geometry and dependency. *Phonetica, 43,* 84-108.

McCarthy, J. J., & Prince, A. (1993). Generalized alignment. In G. Booij and J. van Marle (eds.). *Yearbook of Morphology,* 79-153. Dordrecht: Kluwer.

McCarthy, J. J., & Prince, A. (1994). The emergence of the unmarked: Optimality in prosodic morphology. In M. Gonzalez (Ed.), *Proceedings of the North East Linguistics Society 24* (pp. 333-379). Amherst, MA: GLSA.

McCarthy, J. J., & Prince, A. (1995). Faithfulness and reduplicative identity. In J. Beckman, L. Walsh Dickey, & S. Urbanczyk (Eds.). *Papers in Optimality Theory* (pp. 249-384). University of Massachusetts Occasional Papers in Linguistics 18, Amherst, MA: GLSA.

Michelson, K. (1986). Ghost *r'*s in Onandaga: An autosegmental analysis of **r*-stems. In L. Wetzels & E. Sezer (Eds.), *Studies in Compensatory Lengthening* (pp. 147-166). Dordrecht: Foris.

Miner, K. (1979). Dorsey's Law in Winnebago-Chiwere and Winnebago accent. *International Journal of American Linguistics, 45,* 25-33.

Miner, K. (1989). Winnebago accent: The rest of the data. *Anthropological Linguistics, 31* (3/4), 148-172.

Mithun, M., & H. Basri. (1986). The phonology of Selayarese. *Oceanic Linguistics, 25,* 210-254.

Mudzingwa, C. (2007). Hiatus resolution in Shona. University of British Columbia, Ms.

Munro, P., & P.J. Benson. (1973). Reduplication and rule-ordering in Luiseño. *International Journal of American Linguistics, 39,* 15-21.

Nespor, M. & Vogel, I. (1979). The syntax of word-initial germination in Italian. *Language, 55* (4), 812-841.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology.* Dordrecht: Foris.

Nevins, A. (2010). *Locality in vowel harmony.* Cambridge, MA: MIT Press.

Nevins, A. & Chitoran, I. (2008). Phonological representations and the variable patterning of glides. *Lingua, 118,* 1979-1997.

Nikièma, E. (1989). Governing segments, constituency and syllabification. Paper presented at the 12th GLOW colloquium, Utrecht.

Nikièma, E. (2003). Defective syllables: The other story of Italian sC(C)-sequences. In J. Kaye & S. Ploch (Eds.), *Living on the edge: 28 papers in honour of Jonathan Kaye* (pp. 365-383). Berlin/New York: Mouton de Gruyter.

Odden, D. (1995). The status of onsetless syllables in Kikerewe. *OSU Working Papers in Linguistics, 47,* 89-110.

Odden, D. (2006). Phonology ex nihilo. Paper presented at the Tromsø Phonology Project Group Meeting.

Odden, D. (in prep.). *Kikerewe-English dictionary.* Available at: http://www.ling.ohio-state.edu/~odden/ (downloaded 02/07/13).

Oostendorp, M. van. (1995). *Vowel quality and phonological projection.* PhD dissertation, Tilburg University.

Oostendorp, M. van. (1999). Schwa in phonological theory. *GLOT International, 3*(5), 3-9.

Orie, O.O., & Bricker, V.R. (2000). Placeless and historical laryngeals in Yucatec Maya. *International Journal of American Linguistics, 66*(3), 283-317.

Osborne, C.R. (1974). *The Tiwi language.* Canberra: Australian Institute for Aboriginal Studies.

Pagliano, C. (2003). *L'épenthèse consonantique en français.* PhD dissertation, Université de Nice Sophia-Antipolis.

Paradis, C. & Prunet, J.-F. (Eds.). (1991). *Phonetics and phonology: The special status of coronals.* San Diego: Academic Press.

Paradis, C., & Prunet, J.-F. (2000). Nasal vowels as two segments: Evidence from borrowings. *Language, 76,* 324-357.

Payne, D.L. (1981). *The phonology and morphology of Axininca Campa.* Dallas: Summer Institute of Linguistics.

Piggott, G. (1991a). Apocope and the licensing of empty-headed syllables. *The Linguistic Review, 8,* 287-318.

Piggott, G. (1991b). Empty onsets: Evidence for the skeleton in prosodic phonology. *McGill Working Papers in Linguistics, 7,* 41-71.

Piggott, G. (1992). Variability in feature dependency: The case of nasality. *Natural Language and Linguistic Theory, 10,* 33-77.

Piggott, G. (1999). At the right edge of words. *The Linguistic Review, 16*(2), 143-185.

Piggott, G. (2003a). Theoretical implications of segment neutrality in nasal harmony. *Phonology, 20,* 375-424.

Piggott, G. (2003b). The phonotactics of a "Prince" language: A case study. In J. Kaye & S. Ploch (Eds.), *Living on the edge: 28 papers in honour of Jonathan Kaye* (pp. 401-425). Berlin/New York: Mouton de Gruyter.

Piggott, G., & Singh, R. (1985). The phonology of epenthetic segments. *The Canadian Journal of Linguistics, 30*(4), 415-451.

Ploch, S. (2003). Metatheoretical problems in phonology with Occam's Razor and non-*ad-hoc*-ness. In J. Kaye & S. Ploch (Eds.), *Living on the edge: 28 papers in honour of Jonathan Kaye* (pp. 149-200). Berlin/New York: Mouton de Gruyter.

Poliquin, G. (2006). Canadian French vowel harmony. PhD dissertation, Harvard University. Available at http://roa.rutgers.edu as ROA 861-0906.

Prince, A. (1984). Phonology with tiers. In M. Aronoff & R.T. Oehrle (Eds.), *Language sound structure* (pp. 234-244). Cambridge, MA: MIT Press.

Prince, A., & Smolensky, P. (1993). Optimality Theory: Constraint interaction in generative grammar. Ms., Rutgers University, New Brunswick, and University of Colorado, Boulder.

Prunet, J.-F. (1986). *Spreading and locality domains in phonology.* PhD dissertation. McGill University. (Published in 1992. New York: Garland Publishing.)

Pulleyblank, D. (1995). Feature geometry and underspecification. In J. Durand & F. Katamba (Eds.), *Frontiers in Phonology: Atoms, Structures, Derivations* (pp. 3-33). London: Longman.

Rehg, K.L., & Sohl, D.G. (1981). *Ponapean reference grammar.* In PALI Language Texts: Micronesia. Honolulu: University Press of Hawaii.

Rennison, J. (1990). On the elements of phonological representations: The evidence from vowel systems and vowel processes. *Folia Linguistica XXIV, 3-4,* 173-244.

Reynolds, W. (1994). *Variation and phonological theory.* PhD dissertation, University of Pennsylvania.

Rice, K. (1996). Default variability: The coronal-velar relationship. *Natural Language and Linguistic Theory, 14,* 493-543.

Rice, K. (2003). On the syllabification of right-edge consonants – evidence from Ahtna (Athapaskan). In J. Kaye & S. Ploch (Eds.), *Living on the edge: 28 papers in honour of Jonathan Kaye* (pp. 429-448). Berlin/New York: Mouton de Gruyter.

Rice, K., & Avery, P. (1991). On the relationship between laterality and coronality. In Paradis & J.-F. Prunet (Eds.), *Phonetics and phonology: The special status of coronals* (pp. 101-124). San Diego: Academic Press.

Rosenthall, S. (1997). The distribution of prevocalic vowels. *Natural Language & Linguistic Theory, 15*(1), 139-180.

Schane, S. (1968). *French phonology and morphology.* Cambridge, MA: MIT Press.

Schane, S. (1984). The fundamentals of particle phonology. *Phonology Yearbook, 1,* 129-155.

Schane, S. (1995). Diphthongization in Particle Phonology. In J.A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 586–608). Oxford: Basil Blackwell.

Selkirk, E. (1972). The phrase phonology of English and French. PhD dissertation MIT.

Selkirk, E. (1982). The syllable. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations, Part 2* (pp. 337-383). Dordrecht: Foris.

Sezer, E. (1986). An autosegmental analysis of compensatory lengthening in Turkish. In L. Wetzels & E. Sezer (Eds), *Studies in compensatory lengthening* (pp. 227-250). Dordrecht: Foris.

Shaw, P. (1980). *Theoretical issues in Dakota phonology and morphology.* New York: Garland Publishing.

Silverman, D. (2011.) Schwa. In M. van Oostendorp, C. Ewen, E. Hume, & K. Rice (Eds.), *Companion to Phonology*. Oxford: Wiley-Blackwell.

Spencer, A. (1986). A non-linear analysis of vowel-zero alternations in Polish. *Journal of Linguistics, 22,* 249-280.

Spencer, A. (1996). *Phonology.* Oxford: Blackwell Publishers.

Stampe, D. (1969). *How I spent my summer vacation.* PhD dissertation, University of Chicago.

Stampe, D. (1979). *A dissertation on natural phonology.* New York: Garland Publishing.

Stevens, A.M. (1968). *Madurese phonology and morphology.* New Haven: American Oriental Society.

Tajul, A.K. (2000). *The phonological word in Standard Malay.* PhD dissertation, University of Newcastle. Accessed 12/08/13 from: http://hdl.handle.net/10443/1119

Tranel, B. (1992). The representation of French final consonants and related issues. In, J. Amastae, G. Goodall, M. Phinney, & M. Montalbetti (eds.), *Linguistic studies in Romance languages* (pp. 53-78). Amsterdam/Philadelphia: John Benjamins.

Tranel, B. (1995a). Current issues in French phonology: Liaison and position theories. In J.A. Goldsmith (ed.). The handbook of phonological theory (pp. 798-816). Oxford: Blackwell.

Tranel, B. (1995b). French final consonants and nonlinear phonology. *Lingua, 95,* 131-167.

Trubetzkoy, N. (1939). *Grundzüge der Phonologie* (English translation: C. Baltaxe, 1969). *Principles of phonology.* Berkeley: University of California Press.

Underhill, R. (1976). *Turkish grammar.* Cambridge, MA: MIT Press.

Vance, T.J. (1987). *An introduction to Japanese phonology.* Albany: State University of New York Press.

Vaux, B., & Wolfe, A. (2009). The appendix. In E. Raimy & C. Cairns (Eds.), *Contemporary views on architecture and representations in phonology* (pp. 101-143). Cambridge, MA: MIT Press.

Vennemann, T. (1972a). On the theory of syllabic phonology. *Linguistische Berichte, 18,* 1-18.

Vennemann, T. (1972b). Phonological uniqueness in natural generative grammar. *Glossa, 6,* 105-61.

Vennemann, T. (1974). Phonological concreteness in natural generative grammar. In R. Shuy & C.-J.N. Bailey (Eds.), *Toward tomorrow's linguistics.* Washington, DC: Georgetown University Press.

Wheeler, M.H. (2006). Catalan. In *The encyclopedia of language and linguistics.*

Wiese, R. (1986). Schwa and the structure of words in German. *Linguistics, 24,* 697-724.

Wiese, R. (2000). *The phonology of German.* Oxford: Oxford University Press.

Wilkinson, R.W. (1974). Tense/lax vowel harmony in Telugu: The influence of derived contrast on rule application. *Linguistic Inquiry, 5*(2), 251- 270.

Yavas, M. (1980). Vowel and consonant harmony in Turkish. *Glossa, 14,* 189-211.

Yip, M. (1993). Cantonese loanword phonology and Optimality Theory. *Journal of East Asian Linguistics, 2,* 261-291.

Yoshida, S. (1991). *Some aspects of governing relations in Japanese.* PhD dissertation, SOAS, University of London.

Zaharini, A. (2005). *The phonology-morphology interface in Malay: An optimality theoretic account.* Canberra: Pacific Linguistics.

Zuraw, K. (2003). Vowel reduction in Palauan reduplicants. In A. Rackowski & N. Richards (Eds.), Proceedings of the Eighth Annual Meeting of the Austronesian Formal Linguistics Association (pp. 385–398). Cambridge: MITWPL #44.

Zygis, M. (2010). Typology of consonantal insertions. *ZAS Papers in Linguistics, 52,* 111-140.