"As in syntax, the typical situation is that as one proceeds,
interesting new problems arise at a faster rate than old ones are solved."
Daniel Kahn (1976: 17)

Preface

This thesis aims to contribute to the understanding of various phonological phenomena in natural language, concentrating on English, especially on word-internal and cross-word consonant lenition. It is written in a theoretical framework which is sadly underrepresented on the present-day international scene of academic activity, and whose merits I hope to show, although throughout this work I try to resist the temptation of a theoretical bias and to pinpoint the shortcomings of the analysis, too. The reader is invited to judge the success of my attempts.

The atmosphere of the whole thesis is largely influenced by the impression I got while writing it: the deeper one delves into a subject (be it a survey of linguistic data or a theoretical analysis), the more questions, rather than answers, pop up. This is faithfully reflected in the quotation above from Daniel Kahn's dissertation, which I have chosen as a motto. Nevertheless, I hope to have found the path towards the solution of a few problems.

There are just too many people I owe gratitute to for guiding me along this path. Obviously, I am indebted to the founding fathers (and mothers) of the theoretical framework of my thesis, especially to Jean Lowenstamm for encouraging me from the very beginning; and to Monik Charette for listening to my ideas, giving advice, and for simply being a friend. I am also grateful to my colleagues at SOAS, London, especially Zoë Toft and Sam Hellmuth, for helping me with the data, and the Eötvös Scholarship of the Hungarian State, without which I could not have made it to SOAS. I thank Miklós Törkenczy and Péter Szigetvári for devoting so much of their precious time to discussing drafts of this thesis and previous papers written in the topic, and for lending me otherwise unavailable papers and books. I am thankful to John Harris and Tobias Scheer for answering my emails continually disturbing them, and to a whole lot of

people for sending me (hard or virtual) copies of their books and papers: (in alphabetical order) Eugeniusz Cyran, Lisa Davidson, Stuart Davis, Carlos Gussenhoven, Bruce Hayes, Tobias Scheer, Delphine Seigneur-Froli, Jennifer L. Smith, Graeme Trousdale, and Bert Vaux. In addition, I owe much to the whole of the LinguistList community: to the organizers for creating this tremendous opportunity to communicate with people, and to the members, whom I do not have the space to list, for answering my queries.

I am grateful to all the people mentioned above (and a few others), but Péter Szigetvári and Tobias Scheer in particular, for letting me use (and abuse) their ideas. Of course, none of them are responsible to any extent for the content of the following thesis.

Table of Contents

Chapter 1: Strict CV phonology: Part 1

This thesis is couched within, and aims to contribute to the improvement of, the framework of Strict CV Phonology, a radical syllabic theory that sprung out of Government Phonology (GP – KLV 1985, KLV 1990, Charette 1991, Harris 1994, etc.) at the inspiration of Lowenstamm (1996). Strict CV phonology, as the name already suggests, claims that syllable structure universally reduces to strictly alternating consonantal and vocalic positions. However, before any modifications to the present state of the theory can be introduced in order to widen its scope, a general description is necessary of what that present state of it is and, very importantly, how I understand it. The following discussion focusses on those aspects of Strict CV phonology which make it both a development of GP and a denial of some of its basic assumptions. But first it needs to be pointed out that, just as usual with (phonological) theories, several interpretations exist, which may partly (or even in crucial respects) deviate from the way it is presented here (cf., e.g., Csides 2000, 2002, 2004, Cyran 2003, Scheer 2004, Szigetvári 1999).[1]

Chapters 1 and 2 are devoted to setting the scene for the rest of the thesis. The present chapter introduces the basics of Standard GP, with special attention paid to which features are and are not shared by Strict CV phonology[2], including those aspects of a revolutionary development called VC Phonology (Dienes and Szigetvári 1999, Dienes 2000, Szigetvári 1999b) which are common to both Strict frameworks (henceforth the term CV/VC phonology will be used to refer to such cases). It investigates the fundamental tenets of Standard GP and the implications thereof, trying to show the extent to which GP is what it claims to be, and to which CV/VC phonology can still be considered its sub-branch, conforming to those central principles. It is

[1] It needs to be emphasised at this point that Strict CV phonology and others treated in this thesis are not the only "non-standard" varieties of GP; for a development taking a different direction, see e.g. Head-Driven Phonology (HDP – Hulst and Ritter 1999).

[2] For how and why Strict CV phonology is a logical consequence of GP's evolution, see Scheer (2004).

structured as follows. We take the main assumptions of Standard GP one by one, investigating and evaluating their status and role in Strict CV phonology: Section 1.1 looks into the claim that GP is the phonological equivalent of GB syntax; Section 1.2 describes the GP/CV/VC conception of prosodic structure; Section 1.3 highlights two GP terms, government and licensing, and the way these terms have been (ab)used by the proponents of the theory; then, in Section 1.4, we examine how empty positions can be dealt with in phonology; Section 1.5 briefly introduces Element Theory; and finally, the GP view of lenition is sketched out in Section 1.6.

1.1 Phonology is not different

The research programme of GP (as the name suggests too) arose from the idea that phonology and syntax are but two manifestations of the same cognitive faculty, and thus it is desirable that their theoretical models utilize the same set of descriptive tools.[3]

Generative syntax underwent drastic changes at the turn of the 1970s and 1980s: the basic tenet that S(urface)-structure is derived from D(eep)-structure through a series of ordered transformations was replaced by Government and Binding (GB) theory (from Chomsky 1981 onwards) – (extrinsic) rule ordering and intermediate representations were excluded from syntactic theory, which became more representational, rather than derivational, in nature.

The question arose immediately whether phonology, tailored in Chomsky and Halle (1968) (SPE) and a host of subsequent publications to suit transformational syntax, should go with the flow. The earliest attempts to do so (e.g., Kaye and Lowenstamm 1986 and Majdi and Michaels 1987) were first strongly objected to, the main counterargument being that phonology *is* different from syntax: Bromberger and

[3] Cf. Part One Chapter 8 of Scheer (2004).

Halle (1989) explicitly state that the extrinsic ordering of more than one rule in more than one step is inevitable in phonology.[4]

Nevertheless, GP (and CV/VC phonology), having sprung off the seminal work of some of those who first proposed that there should be a cross-modular symmetry in grammar (e.g. Jonathan Kaye and Jean Lowenstamm), subscribe to the view that phonological theory is non-derivational[5] in essence: processes are triggered by local sources available in the representation, and they take place freely whenever the conditions on their application are met, i.e., in response to parameterised conditions which are locally present in the environment (the principle of Non-arbitrariness). "[...] no process may be prevented from applying to a string by virtue of its position in a putative ordering relationship", says Kaye (1995: 290). Although most GP analyses operate with just two levels of representation (the input and the output), what is more, Scheer (to appear) claims that phonology applies in one single step, the mechanism does not reject derivation in the form of rules applying in sequence, producing intermediate representations. Harris and Kaye (1990), for instance, break down the processes of London glottalling and New York City tapping into two stages, the output of the first being an intermediate representation serving as the input to the second, and the two are intrinsically ordered (see Section 3.2.3). Nevertheless, the GP conception of derivation renders extrinsic rule ordering, or any arbitrary analytic device, impossible.

In addition, GP claims that the phonological component of grammar conforms to Principles and Parameters Theory: it is composed of universal principles and the language-specific settings of parameters. Most practitioners of Strict CV phonology also subscribe to this tenet, but see, e.g., Cyran (2003) for an alternative view. In Standard GP, one of the basic principles of Universal Grammar is the Projection Principle (1) (also called (Prosodic) Structure Preservation (e.g., Harris 1992a: 366) and counterpart to the corresponding principle(s) in syntax), which rules out any restructuring of existing governing relations including ambisyllabicity (Kahn 1976, Clements and Keyser 1983,

[4] The Bromberger-Halle view of phonology is fiercely attacked, among others, in Coleman (1995), Durand (1995), and Kaye (1990c, 1995).

[5] The extent to which GP can be regarded as non-derivational depends on the definition of "non-derivational". See, e.g., Coleman (1995), who accuses GP of being a derivational framework which is "as unconstrained as the models it seeks to replace" (ibid: 344).

etc.) or total resyllabification (Selkirk 1982, Borowsky 1986, etc.), and ensures that "syllable structure is *by and large* preserved from start to finish." (Kaye 1990a: 139 – emphasis mine)

(1) The Projection Principle

Governing relations are defined at the level of lexical representation and remain constant throughout a phonological derivation.

Notice that according to the Projection Principle governing relations, i.e., prosodic/syllabic structure is underlying in GP (e.g., KLV 1990 in Standard GP, Szigetvári 1999b in CV/VC phonology) although a weaker version is adhered to in Charette (1991) or Brockhaus (1995b: 194-5), where only nuclear positions are associated lexically, and syllabification proceeds from the government relations contracted by skeletal slots.[6]

If we take the Projection Principle seriously, it conforms to the standard interpretation of structure preservation, that a transformation or rule should not be able to create or destroy (prosodic) structure, but only to move material around. This leaves us with only three possible phonological operations: melodic spreading into an already existing prosodic position (e.g. assimilation, hiatus filling)[7], segmental composition motivated by factors other than spreading (e.g. fortition[8]), and segmental decomposition (or delinking) (e.g. lenition – see Section 1.6). Notice that the latter two only obey the Projection Principle if we claim it strictly only refers to prosodic structure. Melodic primes within one segment do not contract governing relations, they seem to delete (or at least get underparsed) in phonologically weak positions, and the operation of inserting

[6] Even Kaye (1995: 293) admits that it is likely that there are certain governing relations which are not lexically defined, e.g., certain types of nuclear interaction responsible for stress, tonal and harmonic effects as well as certain cases of inter-onset government.

[7] As Harris (1994: 167) explains, spreading is not viewed here as a dynamic process but rather as a static instruction specifying the phonetic interpretation of a skeletal position.

[8] In contrast to this view, Harris (1990) insists that even in fortition the local source of an element added to the target segment must be found in a neighbouring position. Thus, fortition is just a type of assimilation.

"ambient" phonetic material also appeared quite early in GP (although the question of the local source immediately arises – see the discussion in Harris 1994: 109).

"By and large" in Kaye's explanation of the Projection Principle most probably refers to the situation when adjacent empty skeletal positions (an empty nuclear position and a following empty onset or nucleus) meet upon the concatenation of morphemes, in which some of the original relations are not preserved. When analysing cross-word tapping in English, for instance, Harris and Kaye (1990) and Harris (1994) delete a word-final (empty) nuclear position to explain how a word-final /t/ comes into contact with a following vowel-initial word. Charette (1991), in her account of the behaviour of French definite articles, eliminates the /ə/ of *le* and the /a/ of *la* before vowel-initial words, the second of which no doubt means deleting prosody *and* melody together. Kaye (1995: 317) also mentions a case from Polish where empty positions straddling a morpheme boundary are "removed from the structure". Therefore, in Standard GP it is generally claimed that in the sequence of a vowelless syllable (i.e. Cv) and an onsetless one (cV) the ("extended"[9]) OCP operates in such a way that the first, empty nucleus is "suppressed"[10] (along with the empty onset to its right) and as a result the two syllables are fused (or "superimposed", as Yoshida 1993 puts it), i.e. $C_1v_1c_2V_2$ ends up as C_1V_2.[11] Obviously, this "reduction"[12] violates the Projection Principle (also suspected by Brockhaus 1995b: 212[13] and explicitly stated in Szigetvári 1999b: 102), since the government relation licensing the empty vowel gets deleted together with the empty vowel, in the same way as the government relation linking the empty onset c_2 to its original nucleus V_2.[14] (For more on government relations, see the next section.)

---

[9] Harris (1994: 213).

[10] Harris (1994: 213).

[11] Cf. Section 5.5.

[12] "An empty Nucleus followed by a pointless Onset are removed from any phonological representation in which they occur" – Gussmann and Kaye (1993: 433).

[13] She also remarks that this is an unusual, "surprising" usage of the OCP, which is traditionally thought to require multiple linking of the same melody rather than the total erasure of some prosody and/or melody.

[14] Moreover, considering that most Standard GP-ists think empty nuclei are not completely empty but filled with the so-called cold vowel (cf. Section 1.4), all the deletions mentioned in this paragraph receive even bruter force.

Even in Strict CV phonology there are operations which may be regarded as instances of violations of the Projection Principle hidden "behind the lines". In Lowenstamm (1999), e.g., an empty CV unit (the boundary-marker – see Section 2.1.1), deserted by melodic material, "will simply wither away" at the end of the derivation (ibid: 163), which can only be interpreted as the deletion of skeletal slots.

In the present thesis, a stance requiring a kind of derivational monotonicity[15] is taken, and the Projection Principle is regarded to allow for governing relations to be added in the course of the derivation. In this respect, we follow Brockhaus (1995b), who concludes that new governing relations *are* supplied on the second cycle (and perhaps later, too), at least during the creation of metrical stucture; but we go against the standard view, represented by, e.g., Kaye (1995: 293), which explicitly asserts that the possibility of building constituent structure within a derivation is also completely denied in GP. Notice that such a strict reading of the Projection Principle cannot be upheld, and it has actually been broken on numerous occasions by exactly those who claim to adhere to it the most rigidly.[16] However, we maintain that changing or deleting governing relations or other aspects of suprasegmental structure do cause violations to the Projection Principle, so such operations are undesired in phonological theory and should be avoided. It will be clear from Section 2.2 onwards that the most significant difference between Strict CV and VC phonology lies in how successfully they handle cross-word phenomena, i.e., the situation when words are concatenated – the major subject of the present thesis; and Chapters 4 and 5 will provide an analysis of phenomena like the ones referred to above which does not resort to deletion, elimination, or structural "reduction" of any kind.

* * *

There are (at least) two serious consequences of the GP/CV/VC view of derivation. First, it becomes a *representational* theory of phonology, since one of the key terms in how processes are triggered is the local source. Thus, the way phonological objects (both prosodic and melodic) are represented crucially circumscribes what is and

[15] Cf. Kálmán (1989).
[16] For an even more fierce criticism of arbitrary operations in Standard GP, see Coleman (1995).

is not a possible phonological phenomenon. Second, GP/CV/VC is definitely an *input-oriented* framework (also pointed out in Cyran 2003), in the sense that the Projection Principle clearly requires derivation to observe the conditions on the underlying representation. "[...] when we nowadays speak of the 'phonology of language X', we are referring primarily to the phonological structure of its lexical entries", says Harris (1994: 270). This is in sharp contrast to other contemporary theories such as Optimality Theory (OT) with principles like the Richness of the Base. These two observations explain why GP is unable to account for certain processes, or at least loses certain generalizations, especially in the case of surface conspiracies, that output-oriented theories handle with such ease. As McCarthy (1991) shows, for instance, r-liaison in Eastern Massachusetts English, involving linking-r and intrusive-r, cannot be described with one single rule in a synchronic analysis unless a constraint-based model is used, where a symmetric r~zero relation can be expressed rather than an /r/ -> zero rule or vice versa.[17]

As the discussion unfolds, it will become clear that it is the view of derivation and the organization of grammar that remains the only link between GP (including Strict CV phonology) and GB syntax, besides the overlap in terminological labels; the exact definition and usage of those terms have been radically modified, partly already in Standard GP, but more extensively in CV/VC phonology.[18] As far as the syntax-phonology relationship is concerned, in recent years non-derivational theories (OT – e.g. Prince and Smolensky 1993, 2004; Declarative Phonology – e.g. Coleman 1995) have gained ground, so the parallelism between syntax and phonology is not even questioned. So much so that OT syntax, drawing on the development of OT phonology, is gradually catching up GB syntax in popularity. Surprisingly enough, a development in OT syntax concerns sentence and phrase structure: in fact, it is a categorical denial of it. Newson (2004), for instance, speaks of "deforestation in syntax", i.e., the replacement of constituent structure by the linear alignment of elements, where syntagmatic/lateral relationships receive the main emphasis – a development following the same steps as

[17] The weaknesses of the GP procedural mechanism (and, most probably, the weaknesses inherent in OT representations too) have led to studies combining GP with OT (e.g. Polgárdi 1998, Rowicka 1999, Blaho 2004).

[18] For a detailed comparison of GB syntax and Standard GP, see Honeybone (1999) and Anderson (2004).

were taken by GP-ists in creating CV/VC phonology (see, e.g., Scheer 2004). Thus, it may be the case that phonology and syntax are not *so* different after all, and with time a point will appear on the horizon where the seemingly parallel developments of the two modules meet.

1.2 Prosodic structure and government

GP has introduced the most striking innovations in prosodic structure, based on a strict definition of the major organizing principle of phonology called government, originally introduced in GB syntax. Although the phonology-syntax symmetry is continually emphasized in Standard GP surveys of suprasegmental structure, it was realized at the very beginning that a major feature of syntactic structure does not apply to phonological constituency: recursivity (e.g. Harris 1994: 153)[19]. In addition, Charette (1989) noted that government in phonology cannot be defined in exclusively structural terms, reference to linear adjacency is also needed.

Since Standard GP's innovations in prosodic constituent structure have been carried over to CV/VC phonology, where they have been extended, and it is also this component of phonological theory where Strict CV/VC deviates from Standard GP the most, this section will turn out to be the most prominent in this chapter. Basically, it will present a gradual "farewell to constituency" (Takahashi's (1993) title), eventually resulting in "syllable unstructure", as Tobias Scheer dubbed Strict CV phonology in an EGG School handout (1999b).

Government, a term borrowed from syntax, originally refers to a dependency relation between two skeletal positions: one is head (governor), the other is dependent (governee), the roles being determined by the segmental make-up of the participants (see Section 1.5). While in syntax government is based on the structural configuration

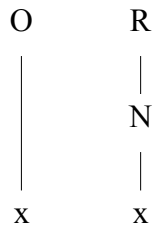
[19] Scheer (2004) takes this as evidence that phonological constituent structure is flat: "the major difference between syntax and phonology is the absence of a tree-building mechanism in the latter module" (ibid: 239).

called c- (or m-) command, already in Standard GP certain forms of government deviate from this definition (see Section 1.3), and, as will be clear in Section 2.1.2, in the theory of Coda Mirror it receives a whole new interpretation. Traditionally, the two essential characteristics of government are (strict or non-strict) locality (i.e., adjacency of head and dependent) and (strict or non-strict) directionality (i.e., fixed right- or left-headedness), the strict interpretations of which logically derive strict binarity (the Binary Theorem – KLV 1990, Kaye 1990a), characterising what is dubbed constituent government and interconstituent government.[20]

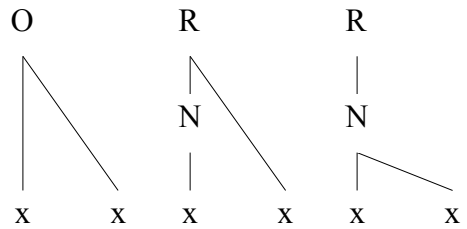
There are three syllabic constituents in Standard GP: the onset, the nucleus, and the rhyme. All of them are maximally binary, which follows from the fact that they are defined as governing domains. Also, all of them may be unary or even null (where a null rhyme obviously means a null nucleus, but a null nucleus may be able to support a rhymal complement, as in, e.g., Magic Licensing – Kaye 1992).

(2) Syllabic constituents in Standard GP

a. Non-branching:



b. Branching:



While empty onsets feature in several phonological theories, the recognition of empty nuclei is one of the distinguishing marks of GP[21]: on the one hand, GP claims that all word-final consonants are onsets rather than codas, followed by Final Empty Nuclei (FENs); on the other, certain surface consonant clusters can only be analysed in GP as sequences of two onsets sandwiching an empty nucleus. The most obvious cases are

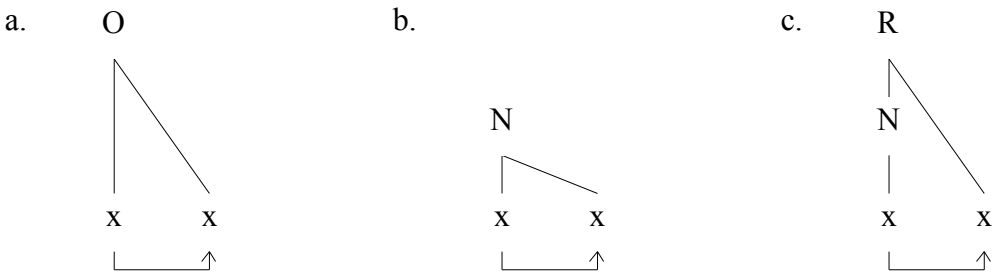
[20] Note that directionality is not governed by parametric variation in these cases, and this constitutes a major deviation from X'-syntax.

[21] Although a number of non-GP-ists have also concluded that empty nuclei do exist, e.g. Anderson (1982), Burzio (1994).

vowel-zero alternations like the Moroccan Arabic examples discussed in Kaye (1990a). Forms like *tan kdib* 'I lie', *tan kidbu:* 'we lie' of the verbal radical |kdb| 'lie', are impossible to analyse without resyllabification (in conformity with the Projection Principle) unless nuclear positions are posited between the consonants. All forms, then, have the structure *kVdVb* underlyingly, with empty nuclear positions separating the consonants; at times these positions are expressed phonetically and at times not, but all such positions are present at the level of lexical representation, and the surface consonant clusters are not true but bogus.

The governing relation that defines the branching constituents is usually referred to as constituent government, and it is head-initial, i.e., it proceeds from left to right.

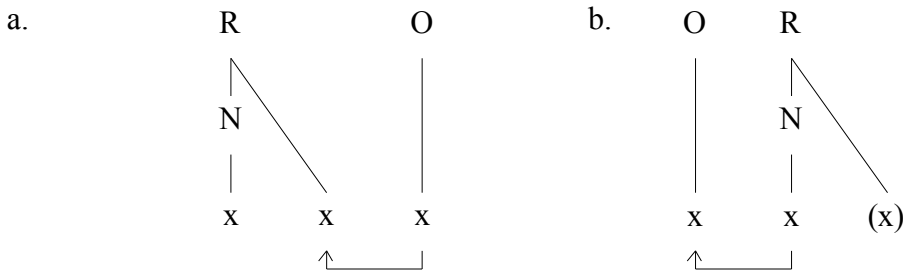
(3)



Apart from empty nuclei, the other distinguishing mark of GP is the rejection of the syllable and the coda as constituents, both on empirical and theory-internal grounds. The main objection to the *syllable* is that it does not participate actively in prosodic (and melodic) processes, and constraints on syllable well-formedness seem to apply to subsyllabic constituents rather than the syllable itself. The theory-internal argument is that were it a constituent, it should be head-initial – but in no way can the onset be considered the head of the syllable.[22] The *coda* is not a constituent either, since it never branches; surface branching codas always coincide with existing coda-onset sequences, and are always head-final (4a).

[22] GP is not alone in denying the status of the syllable in phonological theory: it has also received attacks from other directions (e.g., Aoun 1979, Steriade 1999, McCrary 2002, etc.), although the motivations differ.

(4)



From the onsethood of word-final consonants[23] and the non-branching character of the post-nuclear rhymal complement (= "coda"), it follows that a coda never turns up without a following governing onset. All these facts can be accounted for with reference to Kaye's (1990b) "Coda" Licensing Principle, claiming that "codas" must be licensed by a following onset. This will also derive Onset Maximization (e.g., V.CV rather than VC.V), a principle of syllabification strongly supported by empirical evidence.

The other head-final governing domain consists of an onset and its governing nucleus (4b), and the observation that no onset can exist without a following nucleus is formulated in the Onset Licensing Principle.

A final significant principle is that of government-licensing. Charette (1990, 1991, 1992) observes that consonantal interactions (i.e., branching onsets and coda-onset governing relations) need licence from the following nucleus. Whether (final or non-final) empty nuclei are able to government-license is determined by a language-specific parameter: if it is switched OFF, the language will display otherwise unexpected stable vowels after consonant clusters (e.g. French) or consonant elision (e.g. the Billiri dialect of Tangale).[24] [25]

[23] Also supported by Burzio (1994), who, although working in a completely different framework, analysing the stress system of English, claims that some words end in a vowel. To present the reader with a complete picture, it need be added, though, that this opinion is fiercely attacked in several publications such as Pigott (1999).

[24] A further parameter concerns the difference between government-licensing branching onsets and coda-onset clusters (indirect vs. direct government-licensing, respectively).

[25] Charette's government-licensing principle is reanalysed in Strict CV phonology in Scheer (1997, 1998a, 2004).

Ultimately, then, all constituents are integrated into a network of licensing relations: the Licensing Principle states that all units within a domain must be licensed. The only exception is the head of each domain, which gets licensed at a higher level of phonological projection.

The most exciting type of government, however, is the third: the one contracted by successive nuclei. Most Standard GP models, e.g. KLV (1990), Kaye (1990a), distinguish inter-constituent government between the skeletal slots of successive nuclei from projection government. While nuclear inter-constituent government is strictly binary as strict locality (or adjacency) and directionality (right-to-left) are respected in this relationship, projection government is not strictly binary since it can involve two *or more* nuclei, and is local at the relevant projection only. Moreover, the directionality of projection government is subject to parametric variation (as in X'-theory in syntax)[26]. Nuclear inter-constituent government is contracted by nuclei separated by an empty onset (an empty O without an x on the skeleton), and explains constraints on vowel sequences, vowel raising, etc.; projection government takes place at the level(s) of nuclear projection dealing with prosodic phenomena such as stress systems, vowel harmony, vowel reduction, tonal and syncopation effects.

Proper Government (PG) is a more restrictive case of projection government, and seems to be universally right-headed. PG can only emanate from a non-empty nucleus and attack an empty one, as the result of which it is licensed to remain empty. Kaye (1990a) mentions two possibilities for PG: either it is strictly local (then the result is a binary domain with regularly alternating empty and nonempty nuclei: ØVØVØVØV) or (not strictly) local (then the result is an unbounded domain: ØØØØØØØV), but most examples discussed in the literature are of the binary type.

In CV/VC phonology, "syllable structure universally, i.e. regardless of whether the language is templatic or not, reduces to CV" (Lowenstamm 1996: 419), where C stands for "non-branching onset"/"consonantalness"/"non-syllabicity") and V for "non-branching nucleus"/"vocalicness"/"syllabicity"); syllabic constituency and the skeleton

[26] As e.g. Harris (1994) notes, it is rather frequent to find opposite directionalities at different metrical levels in the same language.

has merged into a so-called CV-tier (cf. (5)), and governing relations are contracted between C and V positions rather than skeletal slots. This is a logical consequence of the introduction of binarity (i.e., the rejection of *n*-ary branching, which is just one step away from the denial of branching altogether) and empty positions (instead of positing empty C's and empty V's between *certain* V's and C's, resp., a maximally constrained theory should have them between *all* occurrences thereof). Once branching is done away with, the distinction between the O-R tier and the skeleton becomes irrelevant unless the need to contrast O's with and without x's, as in Charette's (1991) analysis of French liaison and *h*-aspiré, arises.

(5)[27]

```
closed syllable          geminate                long vowel
     pit                 Hu. ittas 'drunk'          pea
C   V   C   v            C    v   C   V            C   V   c   V
|   |   |                 \       /   |            |   \       /
p   ɪ   t                ...   t       ɒ   ...     p       i
```

The strict CVCV proposal has a number of consequences, some of which Lowenstamm himself mentions in his article:

(i) A language which does <u>not</u> tolerate empty segments will exhibit regular alternances of consonants and vowels; a language which <u>does</u> tolerate empty segments will have apparent consonant clusters and geminate consonants straddling an empty V position as well as long vowels and diphthongs straddling an empty C position. (Cf. Lowenstamm 1996: 420ff.) English and Hungarian clearly belong to the latter type, as illustrated in (5) above.

(ii) Syllable structure for all languages becomes universal.

(iii) Syllable structure for all the words in a language becomes identical. Therefore, the notion of resyllabification loses its relevance (cf. Lowenstamm 1996: 423) and the requirement of prosodic structure preservation, described in Section 1.1, is fulfilled.

[27] I am adopting Szigetvári's (e.g. 1999b) notation which uses upper-case letters for non-empty positions, and lower-case letters for empty ones.

Accordingly, all syllables in the words of a given language have one and the same syllable template, CV, both in lexical and derived representations.

Therefore, the notion *constituent* in the traditional sense becomes irrelevant, since the only "constituents" are the C and the V positions, which, by definition, regularly alternate. Adjacent C and V positions are usually thought of as belonging together, forming "units" (whether they are CV-units as in CV phonology (6a) or VC units as in VC phonology (6b) will be discussed in Section 2.2 and in Chapter 5), but those units are in no way interpreted as constituents.

(6)

```
(a)                                    (b)
c     v     C     V     C     v              v     C     V     C
            |     |     |                          |     |     |
            p     ɪ     n                          p     ɪ     n

            CV phonology                         VC phonology
```

One immediate consequence of such an analysis is the multiplication of empty positions. Members of apparently branching constituents as well as adjacent segments of identical "syllabicity" sandwich empty positions, which are silenced by forces such as Proper Government (PG). (On empty positions, see Section 1.4.) This way, all former consonant clusters become bogus, and the distinction between constituent and interconstituent government, provided it still exists, must be redefined, together with (strict) locality (or adjacency) and directionality. (See Scheer 2004, to appear.) According to one view, since there are no constituents any more, all kinds of consonantal interaction can only belong to the inter-constituent type of Standard GP, i.e., all such relations must be head-final. (See, especially, Scheer 1998b, 2004: 245.) As pointed out by Szigetvári (1999b: 108), the Prosodic Licensing Principle also becomes irrelevant since all skeletal (i.e., C and V) positions are justified simply by virtue of there being a skeleton: V positions are inherently licensed, and CV/VC units are

inseparable – whenever there is a V position it is automatically accompanied by a C position.

Another consequence of CV/VC analyses is the total stripping of constituent structure to the bare minimum: not only do they deny the constituenthood of the syllable and the coda but they reject onsets[28], rhymes[29] and nuclei, too, at least in the traditional sense.

However, CV/VC phonology has a number of advantages over more traditional syllabic models. Lowenstamm (1996) shows how a strict CV skeleton is able to capture the cross-linguistic symmetry characterising the application of processes like compensatory lengthening (as in Standard Italian and Classical Arabic). Adopting a CVCV-template can also insightfully account for non-final closed syllable shortening in languages as diverse as Biblical Aramaic (definitely templatic), Hausa (possibly templatic) and Old English (doubtfully templatic). He turns to Classical Arabic, Chaha, Tiberian Hebrew, Standard Italian, and Danish for examples to justify the claim that there are empty C positions straddled by the members of long vowels, as well as empty V positions straddled by the members of geminates. Larsen (1995) provides a unified analysis of vowel length, Raddoppiamento Sintattico and the selection of the definite article in Modern Italian, using the CVCV model. Szigetvári (1999b, 2000) argues for an analysis of lenition (see 1.6, 2.1.2, 2.1.3 below), "syllable" weight, closed syllable shortening, and compensatory lengthening without the coda constituent and with reference to empty positions. The distinction between heavy and light syllables, crucial in quantity-sensitive stress systems (e.g. Latin, English) and in minimal word effects, for instance, is easily expressable in a uniform and insightful way: light syllables are composed of a single CV span (7a), whereas heavy syllables contain two (7b).

[28] Also attacked independently in Takahashi (1993) and Blevins (1995).

[29] For arguments, see Szigetvári (1999b). The rhyme is also rejected in HDP – see Hulst and Ritter (1999).

(7)

a. light syllable

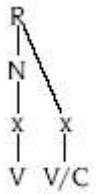
```
C   V
|   |
α   β
```

b. types of heavy syllable

```
C   V   c   V          C   V   C   v
|   |       |          |   |   |
α   β       γ          α   β   γ
```

Compensatory lengthening, as observed in e.g. the two dialectal varieties of the pronunciation of the word *car*, /kɑr/ and /kɑː/, seems impossible to analyse without resyllabification, since the deletion of a coda consonant from a branching rhyme results in a branching nucleus. Even in Standard GP, at least the word-internal examples (such as *card*) involve resyllabification, where post-nuclear rhymal complements turn into the second term of a long vowel. Ritter and Vago (1999) recognize this and reinterpret the "coda" as a dependent rhymal position, which can be either vocalic or consonantal.

(8) The rhyme in Ritter and Vago (1999: 232)



In CV/VC, the same phenomenon is analysed as the underparsing of a consonantal melody due to lack of licensing, and the subsequent spreading of the vocalic content, without any suprasegmental restructuring.

(9)

As shown in Szigetvári (1999b: 35-36), such an analysis also predicts that there are two possible strategies employed in languages: either the vowel preceding the deletion target spreads (as in (9)), or the following segment lengthens. Both directions of change are illustrated in Greek, where 'I am' was pronounced [eːmi] in Classical Attic and [emːi] in Sapphic Lesbian, both originating from reconstructed [esmi].[30]

Hopefully the above discussion has convinced the reader that there are numerous instances where Strict CV models fare much better than traditional theories of prosodic structure: they stress its universal nature, and they do not resort to resyllabification. For even more arguments and examples, s/he is referred to Szigetvári (1999b, 2000), Scheer (2004: 365-744), or Cyran (2003: 205ff). The major issues still posing problems for CV-ers concern the types of consonant clusters (although all clusters are "bogus" in some sense, yet a distinction need be drawn between the types of "bogus clusters"[31] – see Sections 5.3.4 and 5.3.5) and the legalization (i.e., silencing) of empty nuclei.

1.3 What is government (and licensing)?

Both of these two terms, central to GP/CV/VC, originate in syntax, but it has been alluded to above that even Standard GP contains some deviations from syntax as to the

[30] See also the discussion of compensatory lengthening in Scheer (2004: 259ff).

[31] One source of the problem here lies in the fact that there is not even agreement on what exactly those types are. In Standard GP, three kinds of consonant clusters are distinguished (branching onset, coda-onset, and bogus), and much time and space have been devoted in CV/VC phonology to reduce this number to two. For instance, Scheer (e.g. 2004) claims that both traditional coda-onset relations and bogus clusters can be subsumed under the same rubric and analysed as bogus (i.e., with no interaction between the members). Szigetvári (2002a-b), on the other hand, has shown that, at least in English, it is syncope-created bogus clusters that display a behaviour similar to branching onsets.

use of government. For instance, the directionality of government is not a language-specific parameter in most cases but the distinctive feature of constituent vs. interconstituent government. The only exception is (one form of) nuclear projection government, which is claimed to be parametrically either left- or right-headed (see Section 1.2.).

The other term frequently used in Standard GP is licensing, the main force determining phonetic interpretability. Any unlicensed material remains unintegrated into the structure and consequently uninterpreted.[32] Originally government is defined as a special type of licensing (the other type being licensing by parameter setting in the case of final empty nuclei), and that is how empty nuclei may get licensed via PG, or onsets/"codas" receive licence from their governors. Later, in e.g. Harris (1994), government and licensing are more clearly distinguished: government is the force responsible for phonotactic restrictions between certain positions (namely, within branching onsets, branching nuclei, and coda-onset clusters); relations where a phonotactic independence of the participants is found (crucially, the branching rhyme) only qualify as licensing domains but not as governing domains. As you can see, the two terms are mostly, but not completely, interchangeable in Standard GP, so much so that this interchangeability is sometimes rather confusing.

In Strict CV phonology, Larsen (1995), providing a unified account of Tonic Lengthening (TL), Raddoppiamento Sintattico, and the distribution of the forms of the definite article (*il/lo*) in Italian, claims that, governed by a parameter of the phonology, in certain languages such as Italian, long vowels must be properly governed to be licensed. He also posits an empty CV span inserted under stress in TL languages, which must be interpreted segmentally if properly governed. This contradicts the basic definition of PG, that properly governed nuclei remain phonetically uninterpreted. (See also Section 2.2.1 on closed-syllable shortening.)

The inconsistent usage of the expressions "government" and "licensing" eventually led to a kind of terminological confusion, which is seriously criticised in,

[32] For the two subtypes of licensing, viz. a(utosegmental)-licensing and p(rosodic) licensing, see, e.g., Harris (1994).

e.g., Scheer (2004: 149ff), and resulted in new, clear definitions in Coda Mirror (see Section 2.1.2 and Scheer 2004: 211-215).

1.4 Empty positions

Empty categories also belong to the legacy of GB syntax, where several different types of emptiness exist (traces of moved elements as well as static PROs, null Wh-phrases, etc. in addition to positions simply left empty after lexical insertion). The Empty Category Principle (ECP) introduced the notion of Proper Government (PG), required to license traces. GP recognised that empty nuclei require a similar kind of licensing, and therefore the phonological ECP (10) was adopted into the theory. However, in GP a few additional subclauses have proved necessary – only to increase the number of differences from syntax.

(10) The phonological ECP

An empty nuclear position is licensed to remain unpronounced if one of the following holds:

(a) it is properly governed;

(b) it is parametrically licensed domain-finally;

(c) it is enclosed in an onset-to-onset (interonset) governing relation[33];

(d) it is enclosed in an infrasegmental governing relation[34];

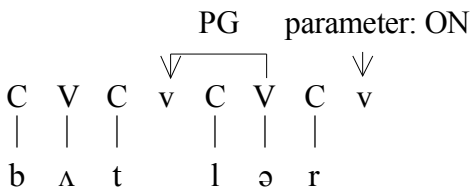
(e) it is magically licensed[35].

Although authors differ as to whether they subscribe to subclauses (c-d-e) of the ECP, the first two are generally accepted in the GP literature. In (11) these main types are illustrated from Strict CV phonology. Via PG, a nonempty nuclear position licenses an

[33] See, e.g., Gussmann and Kaye (1993).
[34] See, e.g., Scheer (1996, 1997, 1999a, 2004: 40ff).
[35] See Kaye (1992).

empty one adjacent on the nuclear (or licenser[36]) projection. The direction of PG is governed by a parameter, but all reported cases illustrate the iambic type.[37] Another parameter circumscribes the presence of domain-final[38] empty nuclei (FENs) in languages: the ON setting (as in English, cf. (11) below) allows for (phonetic) consonant-final words in the language, the OFF setting, on the other hand, implies the absence of such words (as in Italian or Japanese). Kaye (1990b) has shown that the analysis of word-internal codas as true "codas" (i.e., in Standard GP terms, a post-rhymal nuclear complement), and of word-final codas as onsets followed by FENs predicts a four-term language typology (i.e., a language type where both the "coda" and the licensed FEN exist, one where neither exist, and two types with only one of them), which is born out by the facts.[39]

(11)



One of the conditions on PG recognized in syntax is that maximal projections serve as barriers to it. Analogously, in phonology, it has been found that governing domains (e.g., branching onsets, coda-onset sequences) are barriers to PG, i.e., the governor cannot govern across another governing domain. For instance, as Kaye (1990a) illuminates the point, the Moroccan Arabic singular causative form derived from the triliteral root |ktb| 'to write' is [kɨttɨb], where the inter-constituent government between

36 Cf. Charette (1990, 1991, 1992).
37 But see Rowicka (1998, 1999) on trochaic PG.
38 For the difference between "word-final" and "domain-final", see, e.g., Kaye (1995).
39 For a similar four-term language typology related to the distribution of empty C positions, see Chapter 7.

the two members of the geminate consonant, intervening between the two nuclei, prevent the second nucleus from governing the first, producing something like *[kttɨb].

An interesting issue connected to emptiness concerns the content of empty nuclear positions, including FENs. In Standard GP it is generally assumed that they are not completely empty melodically but are latently filled by the cold vowel v^0 (e.g., KLV 1985) or the neutral element @ (e.g., Harris 1994), which is synonymous with the absence of any of the three main vocalic elements (for Element Theory, see the next section). In CV/VC phonology, empty V positions are considered truly empty, with the unpleasant consequence that some studies (e.g. Larsen 1995 or Scheer 1997) were forced to distinguish alternating and non-alternating vowels by lexical marking.

As far as the phonological ECP goes, CV/VC phonology resembles Standard GP in unanimously keeping its first subclause, i.e., (10a), but subclauses (c-d-e) are only utilized to varying degrees.[40] The parametric licensing of FENs (10b) is only questioned in VC phonology (see Section 2.2.2). The "barriers" restriction on PG mentioned above is replaced in Strict CV by a "minimality" explanation: the empty V sandwiched between two consonants (e.g., between the two members of the geminate in [kɨttɨb]) is always closer to the governor and thus the second nucleus properly governs it; but having done so that second nucleus is unable to govern anything else in the structure, so it does not govern the first nucleus simply because it cannot reach it (Scheer 1998a, 1999a).

Therefore, the radical claim about empty nuclei in CV/VC is that they also exist word-internally *after each consonant*, and not only word-finally and in syncopation effects like <batt`ry> where ` indicates an empty nucleus (cf. Harris 1992a: 374), or other bogus clusters. Empty C positions are the direct consequences of the CV/VC representation of long vowels, and they seem to differ radically from empty V positions in that they do not need special care, i.e., the ECP only refers to empty nuclei.[41] This

[40] There have even been attempts to unify the most possible instances of nuclear licensing under PG (e.g., Scheer 1997, 1998a, 2004).
[41] For a discussion of the cross-linguistic treatment of empty C positions, see Chapter 7.

difference in interpretation is explained in Coda Mirror Plus with a difference in definition (see Section 2.1.3).

In GP/CV/VC, then, syllable structure is not surface-true[42], which makes the theory rather abstract. Note however, that GP is not principally abstract owing to intermediate stages arising during the derivational procedure as in SPE, but because of the representations it uses. GP explicitly asserts that there is no linguistically relevant level of phonetic representation (e.g. Kaye 1995), and only constructs which are motivated on purely phonological grounds should be used in phonology. This thesis subscribes to this view.[43]

1.5 Phonological oppositions are privative: Element Theory

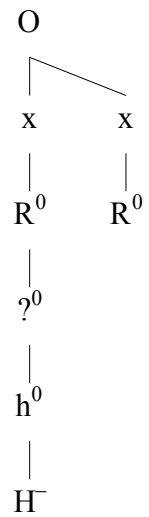
This section introduces the GP theory of segmental structure. Since the present thesis is primarily concerned with prosody and the suprasegmental aspects of melodic alternations, only what is of vital importance for the discussion is sketched out here; consequently, this section only touches upon the topic very briefly.[44]

It has been mentioned in the previous section that in GP terms, there is no level of systematic phonetic representation (Lindsey and Harris 1990, Harris and Lindsey 1993, 1995). Phonological representations are directly interpretable at every linguistically relevant level (the Uniformity Principle – Kaye 1995). Segmental melody is composed of unary (monovalent) primes or elements. Although the exact number and definition of phonological elements varies from author to author (see, e.g., Backley 1993, KLV 1985, Rennison 1998, 1999, Rubin to appear, Scheer 1996, 2004: 40ff, etc.), all versions of GP and CV/VC phonology operate with a relatively small set of unary melodic primes, which are phonetically interpretable both alone and in combination. In

[42] As Scheer (2004: 469) puts it, "What you get is NOT what you see: Tina Turner was wrong."
[43] For a criticism and counterarguments, see Honeybone (1999). In defence of empty positions, see Szigetvári (2003).
[44] For a more detailed yet concise overview, see Appendix 4.1 in Scheer (2004).

the earliest versions of the theory (e.g. KLV 1985), the major characteristic of elements determining their combinatorial possibilities was charm: an arbitrary feature with three possible values (positive, negative, and charmless/neutral). Only elements with neutral charm were free to combine with others; positively and negatively charmed elements "attracted" each other like magnets, elements of identical charm, however, repelled each other. Working with unary primes means there is a difference in complexity between segments, which was found decisive in the case of adjacent charmless segments: in such relationships, the governor cannot be less complex than the governee. Later (Harris 1990, 1994) this "Complexity Condition" was generalized to describe all cases, and the theory of charm, having become redundant, was rejected. For a quick illustration of the point, compare the Standard model of the elemental make-up of /t/ and /r/ in a *tr*-cluster, which can only be analysed as a left-headed governing domain, the first consonant being more complex than the second.

(12) *tr*-[45]



Segmental complexity receives utmost importance in the description of lenition (opening up the possibility of modelling weakening as segmental decomposition – see

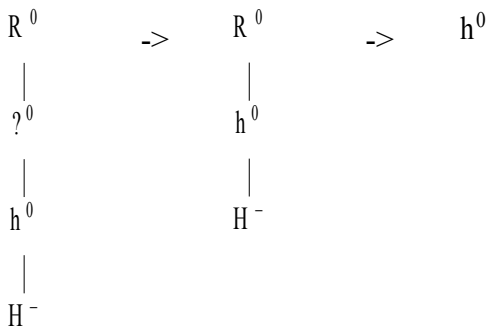
[45] Roughly, **R** = coronality, **?** = occlusion, **h** = noise, **H** = stiff vocal cords. For the sake of illustration, the charm values of the elements are also shown.

the next section), and it may also be conceived of as the reverse of sonority (argued for recently in Cyran 2003).

1.6 Lenition in Standard GP

This is the section where two threads of the GP story meet: the theory describing constituent structure (1.2) and Element Theory (1.5). Recall that in GP/CV/VC a direct causal relation is to be established between a phonological event and its site. Accordingly, lenition (weakening) is considered as element loss due to the weak licensing potential of a skeletal position. Figure (13) illustrates the *t* -> *s* -> *h* lenition trajectory.

(13)



Vowel-zero alternations can be seen as exemplifying the case at issue: a properly governed nucleus is unable to support melody.[46] Although Harris (1992a-b, 1994) builds up a whole theory of lenition (dubbed "Licensing Inheritance", described in Section 3.2.3) based on the intricate system of licensing relations within words, showing how the licensing potential of skeletal positions is depleted gradually in proportion to the "distance" of the position from the head of the domain, it suffers from a number of weaknesses (see, e.g., Szigetvári 1999b, and Sections 3.2.3-3.2.4 below), and so we

[46] Notice the terminological confusion mentioned above: a governed (i.e. licensed) nucleus is unable to license melodic material.

follow a different direction, the one taken by the theory of Coda Mirror and Coda Mirror Plus in Strict CV/VC (Section 3.2.4).

Chapter 2: Strict CV phonology: Part 2

This chapter relates the story of the Coda Mirror (CM – Ségéral and Scheer 1999a) (Sections 2.1.1 and 2.1.2), which ultimately contributed to the birth of VC phonology (or more precisely, the sub-theory of Coda Mirror Plus – henceforth CM+, Section 2.1.3) mentioned above. Since the present thesis draws heavily on both CM and CM+, this story undeniably plays a major role in shaping it. Finally, in Section 2.2 the innovations of VC phonology are confronted with Strict CV phonology.

2.1 CV and the Coda Mirror

This section introduces the advances of Strict CV phonology which are crucially made use of in the present thesis. Just as before, a paper by Jean Lowenstamm starts a stream of theoretical development: his proposal, to represent the beginning of the word as an empty CV unit (henceforth referred to as the boundary-marker) helps Ségéral and Scheer devise a comprehensive theory of lenition and fortition (i.e., the CM), an accompaniment of which is the redefinition of government and licensing; then, as a kind of domino-effect, Dienes and Szigetvári modify CM yielding CM+, and repartition the skeleton yielding VC phonology.

We follow the chronological order of the events this time: Section 2.1.1 introduces Lowenstamm (1999)[1], Section 2.1.2 Ségéral and Scheer (1999a), and Section 2.1.3 Dienes and Szigetvári (1999), Szigetvári (1999b) and Dienes (2000).

[1] See also Chapter 5 in Scheer (2004) for an insightful description of why arbitrary objects like "#" or "pink panther" should be avoided in all scientific activity.

2.1.1 The beginning of the word (1999)

In this influential paper, Lowenstamm describes phonological processes characteristic of the word-initial position but not of word-medial onsets. Since much of the present thesis concentrates on the phonological rules affecting General American (GA) plosives, consider the data of t-aspiration vs. flapping given in (1). They illustrate the point made by Lowenstamm since word-initial t's become aspirated whereas word-medial syllable-initial (single) t's flap[2].

(1)[3] a. [t^h]: Tóm, tomórrow

b. [ɾ]: átom, compétitive[4]

The importance of such phenomena for phonological theory, as Lowenstamm argues, lies in their indication that even the introduction of syllabic constituency to replace linear representations of the SPE-type is unable to banish the word-boundary, #, altogether. Although all t's in (1) are syllable onsets, the phonology treats them differently, which calls for different analyses. In earlier descriptions, phonologists either resorted to ambisyllabicity or rule ordering (see Section 3.2 for more detail). Strict CV Phonology, not having any syllabic constituents, is unable to make reference to ambisyllabicity on the one hand, and, being a theory belonging to the GP family, rejects derivation in the form of a set of ordered rules on the other. The possibility that remains open for Lowenstamm is rediscover word-initial #, which he translates into an empty CV unit.

---

[2] Throughout the chapter, the behaviour of t's appearing in consonant clusters will be ignored since they are beyond the scope of the present discussion. For more detail, see Chapter 3.

[3] The accents indicate the place of primary stress.

[4] For the difference between the two t's in their tendency to flap in words like *competitive*, see Section 3.1.3 and Chapter 6.

(2)

> Rather than being conventionally marked by the insertion of a # symbol to its left, the word is preceded by an empty CV span. The major difference between this proposal and the traditional view lies in the fact that the initial empty CV span is a true phonological site, over which a number of operations will be shown to take place, or in terms of which a number of generalizations will be shown to receive expression. (Lowenstamm, ibid: 157)

Therefore, the representational difference between word-initial and word-medial onsets is that the former but not the latter is preceded by an empty CV unit in the phonological skeleton. This explains, e.g., the presence and absence of certain word-initial consonant clusters in a language like French. To describe the situation in (3), we need two assumptions:

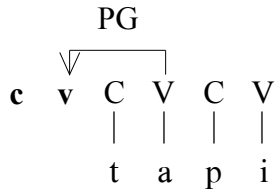
a. In French, the initial site is always licensed (= properly governed).[5]
b. In accordance with Scheer (1996), a sequence of an obstruent and a liquid constitutes a closed domain, transparent for proper government.[6]

In line with this, (3a-b), reconstructed from Lowenstamm (ibid: 159), are well-formed structures in French, whereas in (3c) the initial CV site is not properly governed.

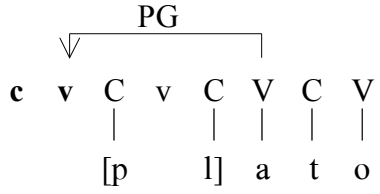
[5] Although it is always the V part of the initial CV that is properly governed, in Lowenstamm's conception of licensing this is enough for the whole CV unit to be licit. Larsen (1995) also speaks of the licensing of whole CV units. It is CM+ where the difference between the licensing/governing conditions on C and V positions is clarified.

[6] The closed domain is brought about by the participating consonants entering into a so-called Infrasegmental Government (IG); see, e.g., Scheer (2004: 40ff) for a detailed discussion.

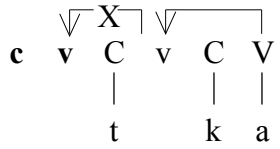
(3)[7] a.French *tapis* [tapi] 'rug'



b. French *plateau* [plato] 'tray'



c. * #tka



A similar analysis accounts for the behaviour of clitics, which, being irrelevant for the present discussion, will not be illustrated here. Still the question remains what accounts for the difference of the t̲'s in the English data in (1). An answer is provided by the theory of Coda Mirror.

2.1.2 Ségéral and Scheer (1999a): Government and licensing as two antagonistic forces

The big innovation came in Ségéral and Scheer's (1999a) new conception of government and licensing[8]. First, the two are opposing relations originating from nonempty V

[7] Single arrows link the source and target of *government*; arrows crossed with an X indicate the impossibility of the relation. Lower case c's and v's are empty positions, while the initial empty CV unit is boldfaced.

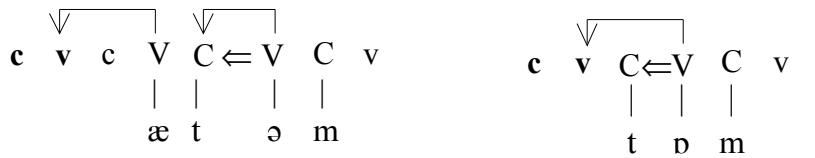
[8] See also Scheer (2004: 134ff, 665-689).

positions in such a way that the same vowel position is capable of contracting both at the same time. Second, they operate between a V position and a V or C position of the strict CV skeleton with the effects specified in (4). Most importantly, both V and C positions can be affected by them, and both in the same way[9].

(4) a. Proper Government inhibits segmental expression of its target.
b. Licensing comforts segmental expression of its target.
(Ségéral and Scheer 1999a: 20)

The /t/ in *atom*, then, is governed (and also licensed) by the following V position, causing it to lenite (5a), as opposed to the one in *Tom*, which is licensed only, the governing potential of the following vowel being consumed by the empty v position of the boundary-marker (5b).

(5) a. *atom* b. *Tom*



The most important message conveyed by the paper is that not only is it possible to subsume the _{#, C} environment of linear models under the single label "syllable coda" (as shown by Kahn 1976; this is equivalent to "a C position *before* an empty vowel" in the Strict CV framework), but its mirror image, {#, C}_, the "Coda Mirror", is also a theoretically relevant site (and translates as "a C position *after* an empty vowel"). In addition, notice that contrary to Standard GP, Onset Licensing becomes optional in this framework.

[9] The idea that (nonempty) V's are able to govern either the preceding Onset position or a preceding vowel but not both is already present in Charette's (1991: 91ff) analysis of *h*-aspiré in French: in *la ache*, *ache* starts with an empty onset which needs PG, therefore its first pronounced vowel cannot properly govern, i.e., silence, the vowel of *la*.

2.1.3 Dienes and Szigetvári (1999): Coda Mirror Plus

The first comprehensive account of English stop allophones in CV/VC was provided in Dienes and Szigetvári (1999), Szigetvári (1999b) and Dienes (2000) under the name of Coda Mirror Plus (proposing an adaptation of Ségéral and Scheer's theory) and VC Phonology (suggesting a skeleton comprising strictly alternating VC units rather than CV spans). For the time being, the division of the skeleton into units will be ignored; the point made by the two authors is the distinction between two types of lenition on the one hand, and the stress sensitivity of lenition in English, illustrated by the word pair *á[ɾ] om- a[tʰ]ómic*, on the other. The basic principles of their theory that are relevant for the present discussion are given in (6) and (7).

(6)

a. Vocalicness is loud, not only acoustically but also in the sense that V slots in the phonological skeleton aim at being pronounced. (Szigetvári 1999b: 62)
b. Consonantalness is mute, if nothing intervenes a C position will stay silent. (Szigetvári 1999b: 62)
c. Government spoils the inherent properties of its target. (Szigetvári 1999b: 66)
d. Licensing comforts segmental expression of its target. (Ségéral and Scheer 1999a: 20)

(7) The Antipenetration Constraint

Government cannot penetrate a stress domain. (Szigetvári 1999b: 79)

The principles in (6) specify the inherent properties of skeletal positions (a-b), which are affected by government (c) and licensing (d) in ways familiar from Coda Mirror. The Antipenetration Constraint (7) ultimately expresses the difference between stressed and unstressed vowels: since the former start a stress domain in VC phonology, they are prevented from emitting government. Thus, in English at least, the /t/ in *atómic* is

treated differently from the one in *átom*: the former escapes being governed and gets aspirated instead (cf. (9a-b) below).

To sum up, a /t/ is aspirated in a phonologically strong position, viz. when licensed but ungoverned; this situation emerges before stressed vowels (since, in accordance with the Antipenetration Constraint, they are unable to govern into a preceding stress domain) and word-initially (when the vowel's governing potential is used up by the requirement to silence the empty V in the boundary marker). There are two types of phonologically weak positions, one is before an empty V, which is roughly before a consonant and word-finally[10] (recall _ {C, #}) – in such cases consonants remain ungoverned and unlicensed and exhibit 'consonantal' lenition (they devoice, deaspirate, debuccalize, etc.), i.e. t̲'s are glottalised. The other weak position is that of foot-internal intervocalic C's, which receive both government and licensing from the following (unstressed) vowel; here consonants tend to move towards vocalicness (they become more sonorous), e.g. GA t̲'s are flapped.

A summary of all the possible consonantal positions is given in (8), taken from Szigetvári (1999b). Since Szigetvári's table does not only cover the cases under investigation here, let us focus our attention on the word-initial and pretonic environments in (8a), and the foot-internal position in (8c).

(8) Possible consonantal positions (Szigetvári 1999b: 135, chart (95)[11])

	LIC'D	GOV'D	LENITION TYPE	POSITION
a.	yes	no	none	#_, oc1, bc2, cc2, _**V**
b.	no	no	c-lenition	_#, bc1
c.	yes	yes	v-lenition	V_*V*
d.	no	yes	c-/v-lenition	cc1, within a long V

The situation in (8c) is exemplified by *atom* in (9a). In both (9b) and (c) the /t/ finds itself in a strong position (subcases of (8a)). Since the skeleton here is partitioned into VC units, stress domains (indicated by brackets) start with the stressed vowels and

[10] In VC Phonology, word-final C's are unlicensed and ungoverned because they are followed by nothing.
[11] Abbreviations: bc*n*, cc*n* and oc*n* mean the *n*th position in a bogus, coda and onset cluster, respectively; **V** is a stressed, *V* is an unstressed vowel.

exclude any consonant(s) preceding it. A characteristic feature of the VC framework is that only consonant-initial words contain a boundary-marker to the left of the first segment.

(9) a. (*atom*) b. (*at*)(*omic*) c. (*vTom*)

```
V  C⇐V  C        V  C⇐V  C  V  C        V  C⇐V  C
|  |   |  |        |  |   |  |  |        |  |   |
æ  t   ə  m        ə  t   ɒ  m  ɪ  k        t   ɒ  m
```

2.2 CV or VC?

We have seen above that the crucial difference between CV and VC phonology lies in how they partition the skeleton. This section enumerates those points where the predictions that the two models make may possibly clash. Let us remark here, though, that the comparison is quite difficult: CV Phonology is very much like a mosaic, with several authors discussing different issues (the most comprehensive version being the one in Scheer 2004 and to appear), whereas VC Phonology is a lot more coherent and elaborate (not being practised by such a number of phonologists).

The difference, then, is whether the skeleton is composed of CV or VC units, neither of which is defined as a constituent in the traditional sense, although Szigetvári and Dienes view VC units as *inseparable* atomic structures (cf. e.g. Szigetvári 1999b: 108). Later, in Chapter 5, this view will turn out to be problematic.

In the following discussion, the implications and predictions of the two theories are confronted in two installments: first, the word-medial situation is under scrutiny, and then we look at word edges and cross-word effects. As it will turn out, word-internally the possible contrasts stem from the way certain principles are expressed (the phonological "software"[12]) (e.g., the Antipenetration Constraint – (7) above) which may heavily rely on the identity of skeletal units and where the boundaries of those units are.

[12] The "software-hardware" distinction is due to Tobias Scheer.

We will see that principles and constraints *can* be stated differently, so that they fit a CVCV skeleton. At word boundaries, however, representation, the phonological "hardware", manifests itself, and issues of the edges (whether they are marked by independent, "mobile" skeletal slots (cv) or by so-called peripheral units like vC or Vc containing melodically filled positions too), and of markedness (whether it is marked to indicate word edges phonologically; what the unmarked syllable template is; etc.) arise. In addition, cross-word phenomena are already mentioned briefly, to illustrate their relevance to the present discussion, but a fuller description is provided in Chapters 4 and 5. Bare in mind that, within the given framework, both aspects of CV versus VC (viz., inside and edge) are crucial since, recall, GP/CV/VC is a representational theory (Section 1.1).

2.2.1 Constraints

The English system of consonant lenition is stress-sensitive, i.e., the intervocalic environment, a frequent site of 'vocalic' lenition (whereby consonants become more vowel-like via sonorization, voicing, fricativization, etc.), which is homogeneous in stress-insensitive languages like Romance, must be broken down into two subcases. It only favours lenition when the vowel following the target consonant is (completely) unstressed; otherwise it is indeed a strong phonological position, resembling the Coda Mirror in hosting stable or strengthening segments. Being a stress-sensitive vs. insensitive system appears to be a parametric choice in languages, unrelated to any other settings (e.g. the strict/permissive dichotomy introduced in Chapter 4). Our theory is expected to be able to express the role of stress accordingly.

In VC phonology, stress domains start with the stressed V to the exclusion of the preceding C position. Therefore the C in a stressed CV sequence is licensed by the vowel but the Antipenetration Constraint, given in (7) above but repeated here for convenience, prevents the same vowel's government charge from reaching it, making it a strong phonological position.

(10) The Antipenetration Constraint

Government cannot penetrate a stress domain. (Szigetvári 1999b: 79)

Although (10) is simply a stipulation, and it does not seem to follow from anything why government, as opposed to licensing, should be unable to enter a separate stress domain, it proves quite useful (as Szigetvári and Dienes 1999 discover themselves). It does not only account for the absence of prestress lenition in English, but also for the absence of prestress syncope (at least word-medially), the PG silencing vocalic positions being just a special form of the government (10) refers to. Moreover, (10) predicts that there are no bogus clusters before stressed vowels, since Szigetvári and Dienes analyse such clusters as not contracting an interconsonantal relation of any kind, the intervening empty v requiring PG from a following nonempty V. Indeed, clusters which are licit word-internally only but not word-initially or finally are considerably rare before stress (cf. *Atlántic*, *ignition*). (10) also implies that different hiatus fillers are expected to appear before stressed and unstressed vowels, since the c in a V_1cV_2 sequence is licensed *and* governed when V_2 is unstressed, but only licensed when it is stressed. This is not borne out it English but Szigetvári (1999b: 134) mentions Malay as displaying such a contrast.[13]

The Antipenetration Constraint, therefore, appears to be a well-grounded one, although it is to be interpreted as a parameter rather than an absolute principle, since it does not apply to stress-insensitive systems. In CV phonology, a possible equivalent would express the ability/inability of stressed vowels to govern into non-peripheral units (see Section 5.2), which would be not less stipulative but would cover the same set of observations about English. A crucial question concerns other stress-sensitive systems (e.g. other Germanic languages): are they characterized by the stress-sensitivity of consonant lenition only, or by the rest of the predictions of (10) too? This calls for some empirical investigation, which is beyond the scope of the present discussion.

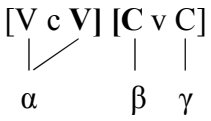
[13] An exciting question that immediately arises is whether Malay can be considered stress-sensitive in all the other respects (lenition, syncope, bogus clusters).

Szigetvári (1999b: Section 7.2.5) argues that closed syllable shortening can be analysed with more insight using a VCVC skeleton. In CV/VC, long vowels are considered to have the underlying structure V_1cv_2, where V_1 is lexically occupied by the melody of the vowel, but v_2 must satisfy certain structural conditions to become an eligible target of spreading for V_1's melody. In the earliest versions of CV phonology (e.g., Lowenstamm 1996, Larsen 1995), that structural condition was thought to be PG emanating from a nonempty V (or a FEN in languages where it is parametrically enabled to (properly) govern) following v_2. Accordingly, closed-syllable shortening characterizes those configurations in which such a governor is unavailable, i.e., when a following consonant cluster sandwiches an *empty* vowel not endowed with governing capabilities. This analysis, however, suffers from several flaws, as pointed out by Szigetvári; for instance, it wrongly predicts shortening before just any consonant cluster other than "branching onsets" (which, following Scheer 1996, form a closed domain and are transparent to PG). Crucially, no shortening takes place before (syncope-created) bogus clusters (cf. *bak<u>e</u>ry, fav<u>ou</u>rite*, etc. with the underlined vowels syncopated), in English at least. Moreover, the CV analysis blurs the functions of PG: it licenses a position to get pronounced when it is the second half of a long vowel, but it licenses a position to remain empty/silent otherwise (cf. Section 1.3). Consequently, VC phonology offers an alternative, with reference to the No Share (or Burial) Constraint.

(11) The Burial (or No Share) Constraint
Burial domains may not share a skeletal unit. (Szigetvári 1999b: 73)

Its effect is that sequences of burial domains (roughly, long vowels, whose members "bury" an empty c, and "coda-onset" clusters, which enter into a C-to-C governing relation – cf. Szigetvári (1999b: 75) – and therefore "bury" a straddled empty v) sharing a VC span are ill-formed. This involves, among others, the situation in question here, illustrated in (12) (on the basis of Szigetvári 1999b: 148, figure 103): a configuration consisting of two burial domains (in square brackets) sharing a VC span (boldfaced) is out.

(12)



This explanation only works for a VC skeleton, and, as Szigetvári himself admits, has no account of the parametric distinction between languages that only display shortening word-internally (e.g. English), and languages that have it word-finally too (e.g. Turkish or Yawelmani) (cf. Kaye 1990b) – a possibility open in CV frameworks able to associate the FEN with various governing and licensing capacities. In Scheer (2004: 253ff), for example, the structural condition required for the vocalic melody to spread is licensing rather than government from a following vowel; all nonempty vowels can license, as opposed to FENs, whose licensing potential is determined by parameter[14]. Although this analysis is not without problems[15], either, at least it demonstrates that CV phonology is also able to capture the phenomenon.

2.2.2 Edges

Originally, the motivation for Dienes and Szigetvári to repartition the skeleton was the realisation that the CV tier contains two "superfluous" locations (see, e.g., Szigetvári 1999b: 90-95): firstly, the initial empty c position, which does not appear to bear any function – the only reason to have it is the simple choice of CV units; whenever an empty v is needed, e.g. to capture the parallelism between word-initial and post-consonantal position, an empty c automatically appears together with it. Secondly, the final v, successor to the FEN, has never been without problems either – recall that it has

[14] For a full typology of the parametric setting of FENs, see Scheer (2004: 651).

[15] E.g., it still does not distinguish bogus and coda clusters (in fact, Scheer (2004) never does throughout the book). Also, if V-to-V licensing is the same force as V-to-C licensing as in CM, then this analysis predicts a difference between intervocalic consonants preceded by short and long vowels, since the first are governed *and* licensed by the following vowel, the latter are governed only, licensing being consumed by the spreading melody of the long vowel. See Section 5.3.3.

received attacks from phonologists of all theoretical tastes other than GP. Moreover, the FEN parameter, whether domain-final empty vocalic positions, that is, consonant-final words, are allowed for in a given language, can be expressed with no less elegance. The result, then, is a VC tier, without the redundant and troublesome empty positions. In CV all words start with (at least) a cv boundary-marker[16], consonant-final words end with a peripheral unit[17] composed of a C and a v, and there is no edge marking in vowel-final words. In VC phonology, on the other hand, two peripheral units are recognized, initial vC and final Vc, but the boundary is not marked in vowel-initial and consonant-final words.

CV phonology claims that empty skeletal positions are dispreferred cross-linguistically, so the unmarked case is when all slots in the skeleton are filled with melody. The exception is the boundary-marker, whose presence may either be universal or governed by a parameter (see Chapter 4), but in neither case is it marked. Therefore, in CV, the unmarked skeleton has the schema cv(CV)*. In contrast, VC phonology has the view that languages prefer to indicate the edges of the word structurally so the unmarked skeleton does not only contain empty positions indicating the beginning but also an empty c to mark the end of the word: vC(VC)*Vc, although in non-peripheral units emptiness is marked here too. Notice that in connected speech the two types of skeleton result in the same string, always a c and a v between the words (provided the boundary-marker is always there – cf. Chapters 4 and 5), so ultimately word-divisions are indicated in an identical manner; however, VC phonology is unable to express the distinction between word-final and utterance-final position in the case of vowels, both being followed by a c, whereas CV phonology is unable to make the same distinction for final consonants, both types being followed by a v. The difference is, should there be need to tell word-final and utterance-final positions apart, in CV phonology this *can* be done by associating different abilities to the two different v's, in the same way as FENs and internal empty vowels are distinguished. If there are no FENs, as in VC phonology, no parametric difference in behaviour can be expressed; and this appears to be the major

[16] Vowel-initial words start with one more empty position, i.e., cvc.

[17] A peripheral unit is a CV-span situated at the left or right edge of the word, containing an empty position.

argument for the existence of final empty vocalic positions, as emphasized quite recently by Scheer (2004)[18].

There are two areas where VC phonology seems to be justified. Firstly, the so-called minimal word phenomenon, widely attested in languages, is much simpler to express (cf. Section 5.3 in Szigetvári 1999b). For languages having such a well-formedness constraint on lexical content words, CV phonology has to say that such words are required to consist of at least two (together with the boundary-marker, three) CV units: (cv)CVCV[19]. In VC phonology, however, a very elegant constraint requiring at least a nonperipheral unit (The Minimal Word Constraint – Szigetvári 1999b: 100) will do the job.

Secondly, vowel-initial suffixes, mostly nonanalytic in English, closely follow the root in VC phonology (...Vc+V...; ...VC+V...) whereas CV phonology suffers from the accumulation of empty positions (...CV+cV...; ...Cv+cV...). Consonant-initial suffixes, on the other hand, are usually analytic in English; unfortunately, it is this case when CV phonology predicts a more intimate relationship between the stem and the suffix (...CV+C...; ...Cv+C...), while on a VC skeleton such suffixes may create cv sequences to mark the morpheme boundary (...Vc+vC...; ...VC+vC...), thus they seem more independent of their stem. (Cf. Szigetvári 1999b: 106.)

The above discussion illustrates that there are almost as many arguments for VC phonology as for CV skeletons. As was shown by Pagliano and Rizzolo (2000), in the analysis of French, CV phonology seems to be more successful in accounting for consonant liaison, but it is unable to treat vowel elision in the case of the clitics *la/le*. For VC phonology, on the other hand, the exact reverse is true.

None of the weaknesses of either framework listed above, however, is fatal: both are capable of expressing, for instance, the minimal word constraint, just one statement is simpler than the other – it is only Occam's razor that would decide in favour of it. As far as the difference between consonant- and vowel-initial suffixes is concerned, a

[18] Another argument in defence of FENs concerns the fact that, by positing a word-final empty category which is licensed/silenced in an unordinary way and which has idiosyncratic properties, CV phonology predicts that the right edge is a special location (Scheer 2004: 71-72).

[19] It is also necessary to add that the first V position is nonempty, to rule out initial clusters, e.g. **tra* is subminimal although it occupies two CV spans.

survey of a much broader scope would be needed to justify that the same analytic-nonanalytic bias is attested in all, or at least in the majority of languages. The major topic of the present thesis, connected speech, nevertheless presents unsolvable problems for VC phonology, as is sketched out presently but elaborated on in Chapter 5. As to the difference between the status of the consonant in a V#CV string and a VC#V string, while both frameworks can explain why a word-initial consonant is always in a strong position, at least in languages like English (Vc#v<u>C</u>V in VC and V#cv<u>C</u>v in CV – the target consonant is protected by a preceding boundary-marker, cf. Chapter 4), the two theories make opposing predictions about word-final consonants followed by a vowel-initial word. A VC skeleton represents the situation in the same way as the corresponding word-internal configuration (V<u>C</u>#V) whereas in CV the boundary-marker is there to separate the two words (V<u>C</u>v#cvcV). Chapter 4 draws on data from several languages to support the claim that there is no phonological resyllabification across word boundaries, i.e., word-final consonants never take up the characteristics of onsets in postlexical phonology. This can only be represented using a skeletal tier composed of CV units. Therefore, we conclude that CV phonology suits our purposes better, and therefore that theoretical framework is chosen in the rest of the thesis.

Chapter 3: The English Cross-Word Puzzle

This chapter introduces the English data the thesis aims to analyse within the theoretical framework outlined in Chapters 1-2. More specifically, it explains what I mean by "the English cross-word puzzle": the behaviour of word-final t̲'s, (pre)glottalized[1] utterance-finally or preconsonantally but tapped/flapped when followed by a vowel. Curiously enough, while word-internally flapping only takes place if that vowel is (completely) unstressed, or else t̲'s surface as aspirated, in the cross-word situation all t̲'s merge into a flap irrespective of stress relations. This story has puzzled theoreticians for quite a long time, and led them to resort to theoretical machineries like ambisyllabification accompanied by rule ordering, resyllabification accompanied by rule ordering, foot construction accompanied by rule ordering, cyclic derivation accompanied by deletion of prosodic structure, etc. All of these previous analyses are rejected in the present thesis for various reasons: if not because one or the other is unable to cover all (relevant) aspects of tapping/aspiration/glottalization, then on the basis of theoretical considerations. As shown in the previous chapter, most of the derivational tools listed above (especially ambisyllabicity, resyllabification and extrinsic rule ordering) are unavailable in GP/CV/VC.

The chapter is organized in the following way: the first part, Section 3.1, attempts to present the relevant data as exhaustively as possible, with special emphasis on the subcases to be analysed in Chapters 5 and 6; and the second part, Section 3.2, sketches out the main trends in previous accounts of the data. Since the phonological literature on English stop allophones is so vast that it is impossible even to dream of providing a full picture of all that has been said on the topic, and since the primary objective of this chapter is rather to give the reader a bird's-eye view of what the thesis rejects and why it rejects it, the section will only highlight the major tendencies rather than supply the complete inventory of previous analyses.

[1] The difference between glottalisation and pre-glottalisation will be clarified in Section 3.1.1.

3.1 The data; or, everything you want to know about flapping vs. aspiration in English

3.1.1 t-allophones: an introduction

As Hoard (1971) observes, there exist a general "tense/lax" opposition in the case of all consonants in English. By "tense" he means a strong allophone, as before a stressed vowel, e.g. the /s/ in *essay*, which differs from the weak, "lax" /s/ in e.g. *messy* at least in length (his and others' acoustic studies have shown that "tense" segments are about twice as long as their "lax" counterparts). This simply illustrates a stress-sensitive lenition process generally affecting English consonants. In certain cases, however, there is more to the variation than just segment length: the allophony of /t/ involves realizations ranging from aspirated through tapped to zero (see (1) below). Also, /t/ seems to exhibit the widest variety of allophony: in many dialects of American English, for instance, it has as many as eight distinct pronunciations, given in (1) (cf. Kenstowicz 1994: 65-66, Wells 1982: 248-252).

(1)	t	"plain"	*stem*
➢	t^h	aspirated	*ten, atomic*
	ʈ	retroflexed	*strip* (only before retroflex /r/)
➢	ɾ	tapped/flapped	*atom, get it in; party, guilty*
	ɾ)	nasal flap	*panty*
➢	$t^ʔ$	unreleased glottalised	*hit, hit me, butler*
	ʔ	glottal stop	*bottle*[2]*, button, sentence*
		zero	*pants*

The variants indicated by ➢ in (1) will have bearing on the cross-word puzzle, so they will be concentrated on in the remainder of this chapter, although reference will from time to time be made to other variants, too. The chapter will also narrow its scope down

[2] Some speakers use a laterally released tap before a syllabic lateral.

to aspiration, tapping and glottalisation as they characterise standard American English pronunciation (General American = GA); the choice as well as potential problems are explained shortly.

The basic facts of GA ṯ-allophony are presented in (2), with the examples taken from Kaisse (1985), Nespor and Vogel (1986), and John Harris (p.c.).

(2)[3]

a.	áṭom at^{h}ómic	Word-internally flapping only applies when the /t/ is followed by an <u>unstressed</u> vowel.
b.	t^{h}én t^{h}omáto at^{h}ómic	Both word-initial and word-internal foot-initial positions are "strong" positions, in which <u>aspiration</u> is attested.
c.	hit^{ʔ}, but^{ʔ}ler hit^{ʔ} me	The glottalised allophone appears before "a consonant and a pause", i.e. in <u>coda</u> position.
d.	hiṭ Ánn hiṭ Aníta waiṭ a mínute aṭ íssue	Across a word boundary the stress-sensitivity of flapping disappears: it also takes place before <u>stressed</u> vowels. Crucially, aspiration does not take place even in that case.
e.	a t^{h}íssue grow t^{h}omátoes	<u>Word-initial aspiration</u> is unchanged in connected speech.
f.	t^{h}o tell the truth t^{h}omorrow[4]	The beginning of <u>grammatical function words</u> is only strong utterance-initially.
g.	I want you ṭo help me Don't lie ṭo me see you ṭomorrow	Utterance-internally lenition affects <u>ṯ</u>'s at the beginning of function words as if there was no intervening boundary.

h. Please waiṭ. I'll be right back.
*They didn't waiṭ. I'll be right back.
Flapping may apply across two words in different sentences but not across just any pair of sentences. Where the two sentences are unrelated, flapping is ruled out.

[3] In the following discussion, two symbols are used to denote the alveolar tap/flap: besides IPA [ɾ], [ṭ] is also applied to avoid interrupting the spelt forms of words and facilitate recognizing them.
[4] Adverbs sometimes behave like function words, cf. (2g).

The aim of the rest of Section 3.1 is to survey the details of the patterning of GA /t/. Any discussion describing English plosive allophony, however, is bound to face a number of problems. On the one hand, the dialectal (and stylistic) variation apparent in previous phonological studies and sociolinguistic surveys forces one to make arbitrary choices to narrow down the subject of investigation. Unfortunately most authors do not state their choices explicitly, hopefully out of forgetfulness rather than the ignorance of the variation; some also oversimplify certain aspects of the data under scrutiny for pedagogic or other purposes; and this shallowness of previous work does not facilitate the understanding of either the data or the analysis of the data. In addition, very often the phonologist comes across phenomena which are scalar on the surface but the linguistic system in question clearly treats them in absolute terms. An example is aspiration, which, as will be discussed in a bit more detail below, appears to be a manifestation of a "continuum in voice onset time" (Kahn 1976: 70) but can still be broken down into a categorical difference between the presence vs. absence of considerable VOT lag; however, where the line is drawn may vary from author to author.

To illustrate dialectal variation briefly, consider the case at issue. In dialects other than GA, t̲-allophony may not apply in an identical manner. Variants not mentioned above also exist, e.g. the spirantization of t̲'s characterizes southern Ireland and the Merseyside area of England both word-finally, as in *get,* and foot-internally, e.g. *letter* (Harris's (1994: 195ff) System D). Other plosives may also undergo this process, and in the case of /t/ it may even be carried on till complete debuccalization to /h/.

In addition, there is considerable dialectal variation in what environments trigger the appearance of the allophones. The extreme case of glottalisation, usually referred to as glottalling or glottal replacement (the debuccalisation of a plosive to a glottal stop), has a restricted distribution in Scotland and most of England (Harris's System A), where it operates utterance-finally only (as in *bit*), but a wide distribution in parts of England (Harris's System B), applying foot-internally too (as in *pity*).

There is also diversity in what segments are affected. In some glottalling dialects, e.g., /p/ and /k/ can be replaced by a glottal stop, too. In the so-called tapping dialects tapping usually affects /d/ as well as /t/ (Harris's System C: most of North

America, Australia, Ireland, and parts of England), and even /n/ (Stampe 1972, Jensen 1987, Marc Picard p.c.).[5]

In sum, any account of English phonological data must state explicitly which English is being described – a fact that has frequently been ignored by authors making statements about "English" by which they mean the accent that they speak. In what follows, we investigate the three-way allophony which /t/ exhibits, and which is to be analysed in the rest of the thesis. GA has aspiration, glottalisation and flapping, whereas RP has aspiration, preglottalization and foot-internal "weakening" (or "laxing" – see above) (cf. Gussenhoven 1986). Of the two standard varieties, GA suits our present purposes better for at least two reasons. On the one hand, it is more frequently analysed in work done within generative phonology; on the other hand, the appearance of the tapped/flapped allophone produces a categorical difference from aspirated /t/, so, even though tapping is optional in certain registers and positions, the difficulties arising from the scalar nature of aspiration (cf. the "twilight zone" in (3) below) can be overcome by looking at less contradictory data. Therefore, after a few introductory remarks on the three-way allophony, the rest of Section 3.1 looks into the phonological environments a /t/ can find itself in.

Aspiration, usually defined as a considerable VOT lag subsequent to the release of stop consonants, devoicing the following sonorant, affects the three voiceless (fortis) plosives and also /tʃ/ (Hammond 1999, Spencer 1996). As hinted at above, aspiration is a multi-valued feature, and on the basis of previous measurements and native intuitions (Kahn 1976: 70, Jensen 1987, 2000: 204, Gussenhoven 1986, Mark Jones p.c., Davis 2005 - referring to an experimental study by Van Dam and Weaver 2001) we can set up the scale in (3). In (3a) example words illustrating the positions unanimously analysed as aspirated are organized in order of degree of aspiration, with ≥ indicating a dubious ordering due to contradictory judgements, while (3c) gives the straightforward cases of no aspiration. (3b), dubbed the "twilight zone", shows the environments in which there is considerable disagreement among the authors. [6]

[5] For all kinds of non-standard variation, see Volumes 2-3 of Wells (1982).

[6] Trousdale (2002) finds a similar scale of phonological positions connected to the increasing likelihood of postlexical ambisyllabification in Tyneside English, which he calls variable ambisyllabicity.

(3) a. aspirated:

tén! > tén ≥ tráin[7] > tèmperaméntal ≥ atténd/discólor ≥ látèx/típtòe[8] > tomórrow > Mèditerránean

b. the twilight zone:

Américan > matrículate > máttress > láter > létter/múppets > áfter > éat

c. unaspirated:

plàstícity > stréss > stém/discúss

The following sections (3.1.2-3.1.6) discuss the tendency of t's to aspirate, tap or glottalize according to phonological position, so I do not go into details here. The foot-initial position, both word-initial (*ten* vs. *tomorrow*) and word-internal (*attend, latex*), emphatically stressed (*ten!*) and with neutral primary stress (*ten*) or secondary stress (*temperamental, latex*), is treated in Section 3.1.2. The foot-internal intervocalic position and the influence of the distance from the stressed vowel (*later* vs. *Mediterranean* and *American*) is under very close scrutiny in Section 3.1.3 and in Chapter 6. Then we turn to preconsonantal plosives and the issue of devoiced sonorants (*train*, *mattress*) in Section 3.1.4, the influence of preceding consonants and its interplay with morphological structure and stress pattern (*after, stress, stem, discolor* vs. *plasticity* vs. *discuss*) in 3.1.5, and the word-final position (*eat, eat it, eat me*) in 3.1.6.

Tapping or flapping is the process whereby a /t/ or a /d/ is replaced by the alveolar tap or flap. Although the question of phonetic precision is not of immediate relevance to the subject of the present thesis, and therefore we will follow the usual practice of using the two terms interchangeably, we may note here that the choice of terminology is not at all trivial. Most authors who do take a stance cite Abercrombie (1967), classifying a tap as a one-tap-trill, and a flap as a gesture involving ballistic movements; according to this distinction, the t-allophone at issue is a *tap* (another authority using this classification is Laver 1994). Ladefoged and Maddieson (1996: 231-232), however, distinguish flaps and taps with reference to the brief contact between the

[7] All sources consulted order *ten* and *train* in this way except for Mark Jones (p.c.).

[8] The order between *temperamental, attend,* and *latex/tiptoe* is difficult to settle due to lack of measurements or comparative studies.

articulators made *in passing* and by a *direct movement*, respectively; therefore American English pronunciation contains *flaps*, as opposed to Spanish, for example. Whatever the authors' phonetic conviction, several different symbols are in use in the literature to denote taps/flaps including /ṭ, D, d, ɾ/, of which we employ two (see footnote 3) (cf. Wells 1982: 249).

As for glottalisation, it affects /p t k tʃ/ in syllable-final position preceded by a vowel (e.g. *bet, watch*), liquid (*belt, hurt*) or nasal (*hint*), /tʃ/ even intervocalically as in *teacher* or *watch it*. In phonetic implementation, British English and American English differ slightly as to the relative timing of the glottal and supraglottal closures. While British English tends to display preglottalization, i.e. the glottal stop *precedes* /p t k tʃ/ (resulting in [ˀp ˀt ˀk ˀtʃ] - Wells 1982: 260-261, Gussenhoven 1986), American English is predominantly characterised by a simultaneous alveolar and glottal closure in [tˀ] (Kahn 1976, Kenstowicz 1994).

The following discussion is primarily based on Kahn (1976), Wells (1982), Gussenhoven (1986), Harris and Kaye (1990), Jensen (1987, 2000) and Davis (2003, 2005).

3.1.2 Foot-initial position

Without any intended theoretical implications, the section subsumes the absolute word-initial (stressed and unstressed) and word-internal pre-stress positions under the same heading. We can do so since t̲'s are resistant to lenition and thus aspirated in both *tomorrow* and *ten/retain/enter**t**ain* in spite of a (linguistically insignificant and therefore negligible) difference in the degree of aspiration (i.e., word-initial unstressed t̲'s are aspirated, although they are shorter than their stressed counterparts, cf. (3)). The stress degree of the following vowel only affects the *degree* of aspiration, but it is attested in *ten!, ten, t̲emperamental, latex, hesitate* alike. In connected speech, word-initial /t/

remains strong and aspirated, as in *buy tomatoes*, but recall from (3) that the beginning of function words like *to* is only strong utterance-initially.

Relatively strong aspiration only characterizes immediately prevocalic t's. Since /t/ plus sonorant clusters, even those traditionally classified as branching onsets, exhibit more variation, they are treated separately in 3.1.4. In addition, the behaviour of the "notorious" *sC* clusters, being the textbook example of zero aspiration, will turn out to be dependent on several factors and consequently require a separate section (Section 3.1.5).

3.1.3 Foot-internal intervocalic positions

In words like *rapid, racket, atom*, no aspiration is attested (and the /t/ is usually tapped). Even in GA some textbooks report weak aspiration in such environments (e.g. *city, fitting*, etc. in Bronstein 1960: 73-75), which Jensen (2000: 204) interprets as a false impression created by the release of an unaspirated and untapped plosive onto the following vowel[9]. Tapping can take place irrespective of whether a primary or secondary stress is found to the left of the /t/, i.e., in *pretty, water* and *photographic, automatic* alike. For most speakers, the quality of the preceding vowel has no effect on the probability of tapping either, so it is the same even when a glide precedes as in *loiter, shouting*. In a few dialects, however, some people have a tap after lax (short) vowels (e.g. *little*, *better*) but not after tense (long) vowels (e.g. *writer*, *later*), as explained in Ladefoged (1993: 65). Word-initial t's and d's in unstressed function words can also undergo tapping in cross-word intervocalic situations, e.g., *say to you*, *I don't know*; besides, adverbs have a strong tendency to "cliticize" onto the preceding word in frequently used expressions, too, e.g. *see you tomorrow, go together* 'match' with foot-internal intervocalic flaps.

However, the situation becomes a bit more complicated when a syllabic consonant is the following segment: if it is a syllabic lateral (e.g. *battle*) then tapping

[9] This weak aspiration, however, seems to be real in British English, cf. Gimson (1980: 155), Wells (1982: 322), Gussenhoven (1986).

applies and a laterally released tap is produced. Similarly, flapping is possible with syllabic /m/ and /r/ (*bottom, butter*), but there is no tap before syllabic /n/: in e.g. *button* the /t/ is voiceless, either nasally released alveolar ([t^N]) or glottal ([ʔ]) (Wells 1982: 251). Glottalisation is particularly frequent when the /t/ is preceded by a non-syllabic /n/ and followed by a syllabic /n/ as in *sentence*, *mountain, maintenance*, etc. (Kreidler 1989: 112, Wells 1982: 251).

As observed by Withgott (1982) and Jensen (1987), taps may be suppressed in certain positions, e.g. the underlined t's tend to be aspirated rather than flapped, and the other voiceless plosives are aspirated, in *militarístic, sanitisátion, monotonícity, Mediterránean, Winnipesáukee, Navratilóva, abracadábra,* etc. They are contained in the third syllable of word-internal dactylic sequences, i.e., between the other unstressed syllable of their ternary foot to the left and the stressed syllable of the next foot to the right. Although Davis (2003, 2005) claims that tap suppression only takes place in non-final dactyls and final dactyls like *sanity* are excluded from its scope, in Chapter 6 I will argue that there is no significant asymmetry between final and non-final dactylic sequences. My suggestion is that a distinction can be drawn between the "weak" position immediately following the head of the foot and a "semi-weak" position following the weak one, irrespective of where in the word the sequence is situated. I base this claim on evidence coming from various sources of the fact that, although in final semi-weak positions the frequency of tapping is somewhat higher and VOT in case of aspiration is shorter, flapping does not apply in such a uniform fashion as in the preceding weak positions: word frequency and speech style can exert their influence there, much the same way as word-medially. For instance, in *capácity* or *éditor* aspiration is more acceptable than in *átom* or *glítter* (e.g. Kahn 1976: 165 fn.17, Hooper 1976, Selkirk 1982, Kenstowicz 1994: 69, Kreidler 1989: 110-111, Vaux 2002 and references therein). Harris and Kaye (1990: 261) also note that in words with two successive potential lenition sites, e.g. *compétitive*, the second is less prone to glottalize or tap than the first, which can be taken as another manifestation of the weak-semiweak distinction introduced above. Chapter 6 will provide more details, a survey of previous explanations, and a new alternative.

There is another group of words which is worth mentioning here since flapping does not apply in a uniform fashion in them, contrary to expectations, and this is the set of words ending in *-to*. It was already noticed in SPE (pp. 190-191) that in some of such words, e.g. *mótto*, flapping is natural, whereas others, e.g. *véto*, can only contain an aspirated plosive.[10]

An interesting issue connected to postvocalic tapping concerns the length and quality of preceding vowels. It is well-known that despite the neutralization of the original voicing contrast between /t/ and /d/ resulting from tapping, the distinction in vowel length universally attested before voiced vs. voiceless consonants is preserved in word pairs like *writing* and *riding*, with the vowels before /d/-flaps being significantly longer (according to measurements by Patterson and Connine (2001), the average vowel length before /t/-flaps is 95 ms, and 111 ms before /d/-flaps). On the basis of their experimental results, Fox and Terbeek (1977) even conclude that the flap itself is not always voiced, but surprisingly enough flap voicing does not correspond to the voicing of the underlying plosive.[11] [12]

[10] Further examples: aspiration in *grotto, Otto, Plato, capo* but no aspiration in *auto, potato, tomato, motto, Vito, echo, Pluto* (Hoard 1971, Kreidler 1989: 111). This observation led Chomsky and Halle to assume a contrast between final zero-stressed phonetic [ow] and final tertiary-stressed phonetic [ow].

[11] Another example of opaque vowel contrasts before taps is served by certain Canadian dialects of English, where "Canadian Raising", i.e., the realisation of the diphthongs [ay] and [aw] as [ʌy] and [ʌw] before voiceless consonants, applies before /t/-taps but not before /d/-taps.

[12] The section does not discuss cases of morphological conditioning, e.g. flapping before deadjectival –*ism* as in *elitism* vs. aspiration before denominal –*ism* as in *magnetism* (described by, e.g., Don Churma in response to a LinguistList query).

3.1.4 Preconsonantal positions

The basic question of t's in preconsonantal position is whether they are glottalized. When followed by an obstruent or a non-continuant sonorant (including /l/), they display a tendency to glottalize, e.g. *catkin, footpath, heights, hearts, beatnik, chutney, Batman, atlas, cutlass*. A preceding [–cons] segment seems to be a conditioning factor, cf. *casts* without glottalisation. In possible onset clusters, i.e. /tr/ as in e.g. *petrol* or *mattress*, glottalisation seems to characterize British English more than GA[13]; instead, the /t/ is plain and the /r/ is frequently devoiced.

The (partial or complete) devoicing of following sonorants in *ápron, mátron, máttress, ácrid, ácclimate* etc. is usually analysed as reflecting the aspiration of the preceding stop (the VOT lag having spread onto them) e.g. in Selkirk (1982), Anderson and Ewen (1987), Giegerich (1992), Kenstowicz (1994), Laver (1994), among others. Spencer (1996: 212-213), however, argues that approximant devoicing is a different phenomenon, also triggered by fricatives as in *Islip, eye-slip; midwifery, fried; shrimp, mushroom*. This is also supported by Hoard's (1971) observation that even in words like *atlas* the /l/ is sometimes voiceless when the /t/ does not undergo glottalisation – it is highly unlikely to be aspirated in that position, so aspiration can not be a possible source of devoicing there. Still it is puzzling why glottalisation and devoicing appear to be in complementary distribution, also described in Harris and Kaye (1990: 270), where they note the two possible pronunciations of the word *battery*: one without glottalling and with an aspirated /r/, and another with a glottal stop replacing the /t/ and with an unaspirated /r/. In addition, Anderson and Ewen (1987: 57) remark that in some varieties of English liquids are devoiced after any voiceless sounds irrespective of their syllabification, so even in words like *athlete*. This example bears close resemblance to

[13] In British English, some speakers use a glottalised plosive and a devoiced approximant in *petrol, mattress, equal, hopeless,* etc. (Higginbottom 1964, Wells 1982: 260, Gussenhoven 1986), which has been taken as evidence that the plosive is ambisyllabic being glottalized and aspirated at the same time, for instance in Anderson and Ewen (1987: 61) and Giegerich (1992: 221), although they argue that this only applies when the preceding vowel is short (e.g. *Cypriot, petrol, macron*), and no glottalisation is attested in *apron, matron, micro*. The authors' inconsistency suggests dialectal variation.

Hoard's *atlas*, and suggests that Hoard is indeed describing a dialect which provides no evidence concerning syllabification.

The (false) impression of aspiration can also be due to the fact that in many dialects, when /r/ stands after /t/ or /d/ within the same word, it combines with them to produce affricate-like sequences, as in *train, drain, attract*, etc, in which the /r/ *is* devoiced. This affrication[14] of /t/ may be mistaken for aspiration, as Jensen (2000) and John Kingston (p.c.) pointed out: in *train* the /t/ is aspirated and the /r/ is devoiced; in *mattress* the /t/ is unaspirated and the /r/ is (less) devoiced. We will see in the next section that /tr/ sequences behave differently from single aspirated t̲'s in another respect: while single t̲'s are never aspirated after tautosyllabic /s/ (e.g. *st̲ay*), the /r/ in *stray* is not less devoiced than in *mattress*. Unfortunately this argument only works for /tr/ clusters, since other approximants in other *sCC* clusters do not get devoiced, cf. *spray* with a voiced /r/ as opposed to *apron* with a voiceless one. Therefore, since /s/ is undoubtedly able to manipulate the aspiration of a following plosive, and the devoicing of a sonorant seems to hinge upon the presence or absence of /s/ as the first member in a cluster, we conclude that sonorant devoicing in a TR-cluster is in fact indicative of aspiration characterizing the preceding stop.

But then a difficulty arises, as pointed out by Anderson and Ewen (1987: 58). While single intervocalic consonants are generally unaspirated when the following vowel is unstressed (cf. words like *rapid, racket, atom* in Section 3.1.3), plosives occupying the first position in a TR-cluster are always aspirated, no matter what the next vowel is, as long as the TR-cluster is not itself preceded by /s/. It was mentioned above that this does not only apply to /tr/, as in *atrophy, seductress* (Anderson and Ewen's examples), where the affrication process explained in the previous paragraph devoices the /r/ anyway, but also in other clusters as in *apron*. Any analysis of t̲-allophones must have something to say about how stress becomes irrelevant to TR-clusters.

One process whereby t̲'s often become preconsonantal is vowel syncope as in *batt'ry*, *pott'ry, t'rrain*. In certain cases the syncopated form has become lexicalized and the /t/ aspirated as in one meaning of *battery* 'cell(s) for supplying electricity'

[14] In what follows, the term "affrication" is used in this sense.

(mentioned above), but normally the /t/ undergoes glottalisation or glottalling as in *battery* 'the act of battering' or *pottery* (Harris and Kaye 1990: 270, Harris 1994: 222-223); word-initially, it remains aspirated without devoicing the /r/ as in the fast-speech pronunciation of *terrain* (which is not homophonous with *train*). The same applies to cross-word situations: normally word-final t's remain glottalized (in both *hit me* and *hit Roy*, cf. the famous *night rate* vs. *nitrate* word pair), but in lexicalized ex-compounds, e.g. *beetroot*, plosives are aspirated/affricated and the following liquids devoiced (as opposed to, e.g., *courtroom*).[15]

3.1.5 Postconsonantal positions

Although the word-final position will be discussed in the next section, final postconsonantal t's are also mentioned here. No tapping or glottalisation is found in *fist, left, fact, apt, after, custard, chapter, doctor*, that is, following an obstruent. Word-medially, when the /t/ is released, aspiration is a possibility after a plosive, as in *chapter* and *doctor*, although its degree may vary. As Selkirk (1982) notes, for instance, the place of articulation of the preceding consonant makes a difference: *chapter* may aspirate more readily than *Acton*. A preceding velar nasal has a similar effect: *Washington* is aspirated for most speakers.

Word-finally following /r/ and /n/, t's are glottalized and never tapped (*hurt, court, start; hurt me, court shoe, start them; hint, paint; print some, paint mark*). Word-internally and when a vowel-initial word follows, however, they are realized as (retroflex) taps, at least after /r/ (*quarter, starter, forty; hurt it, court of, start up*).[16] Following /n/, tapping is generally reported, as in *winter, plenty; print of, paint it*.[17] Kahn claims that the stressed vowel is nasalised and the nasal itself gets deleted in these cases, therefore this is simply a subcase of intervocalic tapping (also in Wells 1982: 251). He argues that many Americans use two pronunciations of, say, *winter*: with *-nt-*

[15] Also in /dr/ sequences: *bedroom* with affricated /dr/ vs. *headrest* with fully voiced /r/.
[16] /d/ can undergo flapping in this environment, too, e.g. *accordion*.
[17] But /d/-flapping is not possible in e.g. *sender* (Kahn 1976: 165, fn.18).

in careful speech, and with a tap but no /n/ otherwise. Wells (1982: 252) even finds some geographical variation: Southerners tend to delete the /t/, Northerners, especially from the east coast, however, preserve the contrast between *winter* and *winner*, e.g. with a nasal flap in the former and a nasal stop in the latter.

After /l/ glottalisation but no tapping is possible word-finally in *fault, belt, bolt; fault me, belt buckle, bolt down*. When a vowel follows, the /t/ is released and thus several pronunciations exist for words and phrases like *shelter, guilty, revolted, Walter; fault of, belt up, bolt it*. For some, tapping is an option (Selkirk 1972, Wells 1982: 251, Bronstein 1960: 73-75), and it isn't for others (Harris and Kaye 1990, Kreidler 1989: 110). As Kahn (1976) explains, the quality of the /l/ is responsible for the variation: if it has a consonantal articulation then no flapping takes place, if there is a non-consonantal (vocalized) /l/ then flapping applies as is usual intervocalically. In addition, Selkirk (1982) observes that non-flap /t/ may optionally aspirate, but *filter* aspirates more often than *alcohol*.

An interesting subcase of postconsonantal /t/ not usually covered in studies consists in occurrences of unstressed t-initial function words following consonant-final words, e.g. *talk to, have to, been to, sell to, (ask) her to*. Looking at their behaviour we can conclude that the underlined t's surface as they would in the corresponding word-medial position, i.e., *talk to = doctor, have to = after*, etc.

Leslie (1983) and Harris and Kaye (1990: 271) note the anomalous behaviour of the '13-14' set of items (or the *-ee/-oo* cases, as Harris and Kaye call them). They list *thirteen, fourteen, eighteen, nineteen, canteen, frontier, settee; pontoon, cartoon, tattoo, spittoon; seventeen, guarantee* as words in which t's lenite although they are followed by stressed vowels. They report wide-spread glottalling in London vernacular and quite common tapping in tapping dialects including some types of Australian English (though not in New York City). In those items of the list where the /t/ is not intervocalic, the preceding consonant is either /r/ or /n/ - the sonorants of *party* and *winter* where tapping is normal. Crucially, there is no lenition in *fifteen* and *sixteen* in any of the dialects displaying the phenomenon, so whatever triggers the lenition here it has the same segmental conditions as "ordinary" flapping (cf. *after, custard*).

The most exciting postconsonantal /t/, however, is the one preceded by an /s/. In the remainder of the section we will outline the circumstances under which aspiration is prevented in /s/ plus plosive clusters. The question why /s/ kills off aspiration is difficult to answer. First of all, it is not clear whether a phonetic or a phonological explanation is to be found. Phoneticians have frequently referred to some kind of economy of glottal movements (e.g. Kim 1970): the glottis opens during the production of /s/ to anticipate the stop, begins to close by the time the closure of the stop is made, and by the time of the release the glottis is too narrow to generate aspiration. Be it as it may, this is in no way a universal phonetic feature; Classical Greek is well-known for distinguishing between aspirated and unaspirated plosives after /s/. Therefore, the absence of aspiration requires a language-specific account.

Secondly, contrary to general belief, it is not only /s/ that can prevent aspiration, but other voiceless fricatives, too. In GA, for instance, the /t/ is unaspirated in *after* in the same way as in *custard* (Kahn 1976) (in contrast to British English, which Gussenhoven (1986: 124) claims to unquestionably have aspiration in *after* although not in *distill*). In German, /ʃ/ is the fricative most frequently preceding voiceless plosives word-initially, and it has the same effect: no aspiration in *Stab* 'stick', *Spiel* 'game', etc. (e.g. Wiese 1996: 270).

Phonologists usually account for these facts by letting syllabification rules capture the /s/ into the onset, and specifying the absolute syllable-initial position as the site of aspiration – thus excluding plosives preceded by the /s/. This solution receives additional support from the factors generally believed to influence syllabification: stress pattern and morphological structure. Davidsen-Nielsen (1974) finds that while the plosives in *spin*, *despise* are unaspirated, when a word contains a morphologically transparent prefix with *–s,* e.g. *miscalculate*, *discourteous,* the following stop *is* aspirated. This is contrasted to opaque prefixes as in *discard*, where the plosive is just as unaspirated as in *discuss*. He also recorded that secondary stress on the preceding syllable as in *gèstátion, fàstídious* contributes to a somewhat greater average degree of aspiration than in *bestow* and *establish* with no initial stress.

Jensen (1987) unites the two cases by proposing that transparent *mis*- and *dis*- exhibit secondary stress, and carries out a follow-up experiment to test Davidsen-Nielsen further: he concludes that *ìnfèstátion, èlàstícity, plàstícity, pèstíferous, òsténsible* have stress on the preceding syllable, and consequently a greater duration of the release stage of the plosive, i.e., more aspiration. In contrast, *askánce, órchestra, astónish, pédestal, sustáin* have no initial stress, and a shorter release phase.

sC clusters created by word-initial schwa-deletion provide another example of aspirated consonants following /s/. Apart from clearly lexicalized forms like *s'ppose* for *suppose*, and for some speakers, a few other words too e.g. *s'pport* for *support*, the original aspiration is preserved and as a consequence, word pairs like *spear* and *S'pir* for *Sapir* are not homophonous.[18]

A related issue was touched upon in the previous section: the voicing of sonorants in *sCC* clusters. Recall that the realisation of /l, j, w/ following an s-stop cluster (e.g. *splash, steward, squeal*) is almost fully voiced, the devoicing in *str*- is most probably due to the affrication process described above. Since affrication is a completely different process, it is not surprising that a preceding /s/ will not influence it. In the other cases there is no affrication and of course there is no (extensive) aspiration either, and consequently no (extensive) devoicing of the sonorants.

Finally, let me point out that there may be even more factors responsible for the presence or absence of aspiration in *sC* clusters; Vaux (2002) mentions in passing, e.g., that the /k/ in *Wisconsin* is not aspirated in its Wisconsin pronunciation, but aspirated outside of that state, which suggests that word frequency seems to have a say, too. The higher the frequency of an item, the more likely its *sC* cluster is to be analysed as a syllable-initial sequence, analogously to its word-initial occurrences.

[18] I am grateful to Lev Blumenfeld for sending this and other creative examples.

3.1.6 Word-final position

Postvocalic word-final t̲'s present the core of the cross-word puzzle: utterance-finally they are generally unreleased[19] and often glottalized (e.g. *get*), utterance-medially they exhibit stress-insensitive tapping (i.e., both in *get away* and *get out*).[20] In the latter situation the /t/ can be glottalized if and only if the phrase is spoken with a pause; but glottalisation is mandatory when the following word starts with a consonant (*get me*).

Several authors claim that word-final pre-pausal plosives are (at least weakly) aspirated, mostly for British English (e.g. Wells 1982: 322 for London English), but also for GA (e.g. Vaux 2002); basically, the more emphatic and energetic the speech, the greater the likelihood of aspiration. Laver (1994: 355), however, attributes the impression of final aspiration to the audible release of the voiceless stop. Surprisingly, Borowsky (1986: 268) even allows (optional) flapping word-finally: *ge[D]#, ha[D]#.*

Following consonants before a pause, the release is obligatory if the consonant is an obstruent (*apt, list, clasp, ask* - with final aspiration for Spencer (1996: 209)). When it is /l/, the release is possible (*alp, belt, milk*); when it is a nasal or /r/, the plosive is unreleased (*camp, can't, hank; heart, harp, hark*) (cf. Kahn 1976). Recall from the previous section that all word-final plosives behave as word-internal ones across word boundaries, except that the stressedness or otherwise of a vowel in a following vowel-initial word is irrelevant (i.e., *start up = party, belt up = shelter, paint it = painter*, etc.).

3.2 Previous analyses

This second part of the chapter introduces and evaluates some of the previous analyses of English t̲-allophones. Since what happens melodically is secondary to the present discussion, possible descriptions thereof (e.g., aspiration as a "puff of air" or "a period

[19] Kahn (1976: 81) reports a difference in the naturalness of release between plosives in this position: /t/ must be unreleased, /k/ tends to be released, /p/ is somewhere inbetween.
[20] Compare the American English pronunciation (with a tap) and the British English pronunciation (with aspiration) of *at áll*. While AmE treats it as a phrase, in BrE it behaves as one single word.

of voicelessness", or, as Laver (1994: 37) puts it, "a moment of relative silence") and its theoretical models (e.g., various references to VOT, binary features like [±spread glottis] or [±Tense], or even [extra-short closure] for flaps as in Steriade 2000) are left unaddressed.

In the rest of the chapter, then, we concentrate on how the suprasegmental structure thought to trigger the segmental changes is modelled, and examine what answers some of the earlier analyses give to the questions in (4).

(4)

- What causes the stress-sensitivity of lenition in English word-internally and the disappearance of this stress-sensitivity across words?
- What do word-initial and pre-stress positions share that makes them phonologically so strong that they are protected from lenition in connected speech?
- Why are there different types of lenition in foot-internal intervocalic positions and in codas?
- How are function words different from lexical content words?
- What is the domain of flapping?
- How is the behaviour of /t/ in consonant clusters explained?
- How is the behaviour of /t/ before syllabic consonants explained?

The following discussion intoduces those previous analyses of English stop allophony that have some bearing on the Strict CV account to be offered in Chapter 5: they are either written in a mainstream framework but are flawed enough to justify the search for an alternative, or they have influenced the analysis in Chapter 5 directly by lending it some insight. Since the discussion proceeds roughly in a chronological order, it also reflects the development of generative phonological theory with regard to consonant lenition. Section 3.2.1 describes the ambisyllabicity analyses of Kahn (1976), Gussenhoven (1986), Rubach (1996) – one from each decade of the last third of the 20th

century, hallmarked by the renaissance of phonological constituent structure. Then Section 3.2.2 expresses the resistance to improper bracketing: how Kiparsky (1979), Jensen (2000) and others stand up against ambisyllabicity, and apply resyllabification and rule ordering to derive the same results.[21] This route is followed by most proponents of Prosodic Phonology (PP), too, with Nespor and Vogel (1986) as a defining publication. Section 3.2.3 does away with both resyllabification and ambisyllabicity in the paradigm of Standard GP, and lenition is treated in terms of government and licensing; finally, Section 3.2.4 does away with traditional constituent structure altogether and redefines both government and licensing as done in Ségéral and Scheer (1999a) (CM) in CV phonology and in Coda Mirror Plus (CM+) in VC phonology (Dienes and Szigetvári 1999, Szigetvári 1999b, Dienes 2000). Finally, Section 3.2.5 concludes.

3.2.1 Ambisyllabicity

It was recognized rather early, namely in Daniel Kahn's oft-cited dissertation in 1976, that phonology should restrict its scope to phenomena that are productive enough to be proven to have a kind of "psychological reality" for native speakers. Thus, attention was shifted from SPE's laxing rules to allophonic processes like aspiration, tapping, etc., which are most insightfully described in terms of syllable structure. Already in Kahn (1976), the notion of ambisyllabicity is introduced to account for the tripartite variation displayed by /t/ (viz., aspirated, glottalized, tapped). Ambisyllabicity is a segment's or skeletal position's simultaneous association to the preceding *and* the following syllable, brought about by two mechanisms, Coda Capture (within morphemes) ((5a) below) and Onset Capture (across morphemes) ((5b) below) applying upon the output of core syllabification.[22] Therefore, aspiration is expected when a plosive is syllable-initial and

[21] In contrast to the present survey, Harris (1994: 198ff) treats ambisyllabicity and resyllabification analyses in a single discussion, under the heading of "Coda analyses".

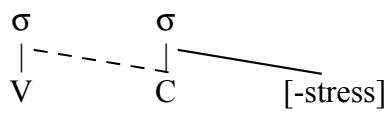
[22] Ambisyllabicity is usually called for to resolve the conflict between the Maximal Onset Principle (MOP) and the Complex Rhyme Condition (CRC) (cf., e.g., Trousdale 2002) in words like *city*: the /t/ here is assigned to the second syllable by the MOP but also required to close the first by the CRC. The

non-syllable final; flapping applies in ambisyllabic position[23]; and glottalisation occurs when a plosive is syllable-final and non-syllable initial.

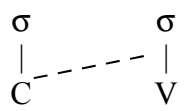
Kahn's approach, the first to analyse American English stop allophones with reference to ambisyllabicity, was later adopted and extended for glottalling in British English dialects by Leslie (1983), and for other aspects of stop allophony by Gussenhoven (1986) and Rubach (1996), among others. Essentially, in all versions of the ambisyllabicity analyses, there are two rules of ambisyllabicity in English: on the one hand, a word-level lexical rule of Coda Ambisyllabicity (Rubach 1996: 221, cf. Kahn's (1976) Rule III, Gussenhoven's (1986) Right Capture), which is stress-sensitive (5a); on the other hand, a postlexical Onset Ambisyllabicity rule (Rubach 1996: 222; cf. Kahn's Rule V, Gussenhoven's Liaison), which is stress-insensitive (5b).

(5)

a.



b.



In (5a), a consonant at the beginning of a syllable headed by an unstressed vowel is associated to the preceding vowel-final syllable. In (5b), a syllable-final consonant is linked to the following onsetless syllable. The two ambisyllabicity rules are followed by the rules spelling out the two lenited allophones of /t/, viz. the tap in ambisyllabic position and the unreleased/glottalized stop syllable-finally, together referred to in Harris (1994: 199ff) as Lenition. As Harris (ibid: 200) points out, the ordering of Coda

other way to resolve the conflict is by means of resyllabification, to be considered in the following section.

[23] In fact, Kahn (1976: 98-104) claims that ambisyllabicity need not be formally incorporated into the flapping rule. He breaks flapping and glottalisation down into two stages: they share the process he refers to as Sonorization in syllable-final position before a [-cons] segment (he considers glottalized /t/ as [+sonorant]), but then Voicing takes place before a [+syll] segment (yielding a tap) or else Glottalization ensues. If Glottalization is ordered after Voicing, which is in turn ordered after Aspiration/Sonorization, the correct results are generated without recourse to ambisyllabicity in the formulation of the rules itself.

Capture (5a) before Lenition is intrinsic since no taps can be produced until ambisyllabic consonants are created, while the ordering is extrinsic between Onset Capture (5b) and Lenition: the opposite order would generate an unreleased stop in the cross-word configuration e.g. *get off*. Due to the latter feature of the analysis, it has not escaped theoretical objections to the effect that it contains a stipulative mechanism.

Turning to the key issues listed under (4) above, we find that the stress-sensitivity of lenition word-internally and the disappearance of this stress-sensitivity across words are directly built into the rules themselves, and do not follow independently from any other principle. The fact that the word-initial and pre-stress positions are both phonologically strong and are protected from lenition in connected speech is ensured by their being associated to one syllable only (rather than being ambisyllabic) and by the absence of a third type of ambisyllabicity, a postlexical version of (5a) where a syllable-initial consonant is attached to the preceding syllable (irrespective of stress relations, like (5b)). The reason for the different types of lenition in foot-internal intervocalic positions and in codas lies in the formulation of the allophony rules: the flap rule applies to ambisyllabic /t/, whereas glottalisation takes place in coda-only position. The ambisyllabicity accounts are silent about how function words differ from lexical content words, what happens before the various syllabic consonants, and what the domain of flapping is.

The behaviour of /t/ in consonant clusters is explained directly by its syllabic affiliation: in TR-clusters it is syllable-initial and therefore aspirated, but it is syllable-final when first in a heterosyllabic cluster (as in *butler*) and glottalized, and it is neither in *sC*-clusters, which are generally syllabified as onset clusters.

In spite of the wide empirical coverage of ambisyllabicity analyses, even in derivational models it is generally acknowledged that rule ordering can produce the same effects without ambisyllabicity, that is, without the improper bracketing that loosens the theory. This is what the next section is about.

3.2.2 "Against ambisyllabicity"

As mentioned above, ambisyllabicity is not widely accepted in phonological theory. Numerous efforts have been made to eliminate the need for it in English phonology, and these efforts are faithfully represented by the title of Jensen (2000), chosen as the title of this section. Already in 1979, Kiparsky attacked the ambisyllabicity model, transferring the generative burden to the derivational machinery in the form of a cyclic analysis, and the hierarchy of phonological constituents in the form of reference to the foot. Thus, in his analysis, the /t/ is laxed (i.e., "weakened") on the word cycle in non-foot-initial position; post-cyclically, this lax /t/ is tapped syllable-initially or glottalized syllable-finally, while nonlax voiceless stops undergo aspiration. To derive cross-word tapping, a phrase-level resyllabification rule applies before the tapping rule proper goes into effect. That is, the complete dissociation of the word-final consonant from its original syllable takes place – two operations in fact: the consonant is first delinked from its position yielded in core syllabification, then it is relinked to another. This is complete Onset Capture.

Selkirk (1982) is written in a similar vein, but with the opposite directionality: under certain stress and segmental conditions, it is a (single) *onset* consonant that is detached and adjoined to the coda of a preceding syllable, i.e., complete Coda Capture takes place. Both tapping and glottalisation occur in syllable-final position, but there is a feature contrast [±release] introduced, where the plus value triggers tapping, i.e., voiceless stops undergo glottalisation when they are unreleased, while tapping affects syllable-final released plosives preceded by a [-cons] segment.

Borowsky (1986) (evaluated in Harris 1994: 200-203) also propagates resyllabification. She redefines Coda Capture to apply within the foot, and provides a two-stage analysis with reference to (Selkirk's) [±release]: plosives are [-release] before a consonant or pause, and [+release] elsewhere. In the second stage, [+release] yields a tap in a coda, [-release] an unreleased glottalized plosive. The advantage of this analysis is that it requires no extrinsic rule ordering: tapping is intrinsically ordered after Coda Capture, and both lenition rules are intrinsically ordered after the rule specifying the

values for [release]. However, as pointed out in Harris (1994: 202), in this intrinsic rule system lenition hinges on the specification of [release], which in turn hinges on syllabification; but [release] appears to be an ad hoc feature, and since both the tap and the unreleased allophone arise in coda, the distinction is made with reference to good old "before C or pause", that is, the undesirable disjunction is back.

In his criticism of resyllabification analyses, Rubach (1996) argues that in cases when the foot-internal intervocalic allophone coincides with the foot-initial allophone (e.g. clear/dark /l/ in RP), we face the Duke of York gambit: core syllabification assigns the intervocalic consonant to an onset, then Coda Resyllabification moves it into a coda to create a heavy (closed) stressed syllable, and then Onset Resyllabification moves it back into an onset to produce the right allophone.[24]

In addition, notice that resyllabification analyses are unable to separate the two basic lenition sites, the coda and the foot-internal intervocalic position, although they favour different segmental processes, not only in English but it is a universal tendency, too. While ambisyllabicity theories distinguish the strong onset from both the coda and the ambisyllabic consonant, accounts making reference to resyllabification must necessarily merge two of them: for Kiparsky, both aspiration and tapping are associated with the onset; for Selkirk and Borowsky, both tapping and glottalisation take place in the coda.

The problems posed for resyllabification analyses cannot only be resolved by turning back to ambisyllabicity, as Rubach (1996) did, though. Prosodic Phonology (PP), a theory of phonological domains, achieves the same objective with the crucial use of the hierarchy of prosodic constituents (consisting of the following units: (mora), (segment), (syllable), (foot), phonological word, (clitic group), phonological phrase, intonational phrase, utterance)[25] and rule ordering.[26] The usual PP derivation (e.g. Nespor and Vogel 1986: 77, 224ff) contains an early tensing/aspiration rule in foot-

[24] The same argument is cited in Gussenhoven (1986: 134).
[25] The phonological units given in parentheses only feature in some of the analyses written within the framework of PP; some of them, e.g. the clitic group, are even rejected as prosodic constituents by most authors.
[26] For further details on PP, see Chapter 4.

initial position, first followed by flapping across the board within the utterance, then by glottalization. As a more recent example, Jensen's (2000) analysis is given in (6).

(6)

(i) Tensing of all consonants in foot-initial position;

(ii) Aspiration of [+tense] voiceless noncontinuants;

(iii) Flapping of [-tense] nonstrident coronal stops between a [-cons] segment and a vowel within the utterance;

(iv) Glottalisation of [-tense] voiceless stops between a sonorant and zero or more syllable-final consonants.

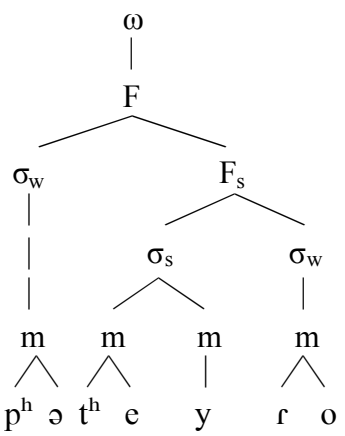
Although this account does not need ambisyllabicity or resyllabification, and it makes use of intrinsic rule ordering only, it still fails to establish a logical connection between prosodic position (i.e., foot-initial, utterance-internal intervocalic, syllable-final, etc.) and its effect (i.e., tensing, aspiration, flapping, glottalisation). In fact, it only makes scarce reference to syllable structure itself: notice that (6iii), for example, is expressed in mere segmental terms. Apparently, representational restrictiveness can only be achieved by sacrificing some of Kahn's original insights.

Nevertheless, PP has made its contribution to the analysis of t-allophony: it has identified the domain of flapping, which is the phonological utterance. As pointed out by Nespor and Vogel (1986: 46), flapping may apply across two words within a sentence, and across two words in different sentences but not across just any pair of sentences. Where the two sentences are unrelated, flapping is ruled out. Therefore, the notion of phonological utterance (*U*) is introduced, and flapping is identified as a *U*-level rule. The phonological utterance is the largest constituent in the prosodic hierarchy. Like the other prosodic constituents, it makes use of syntactic information in its organisation but is not necessarily isomorphic to any syntactic constituent; moreover, its structuring does not only depend on phonological and syntactic factors but is also driven by factors of a logico-semantic type. The crucial factor determining whether two or more sentences form a single utterance seems to be the nature of the relationship

between the sentences. Namely, adjacent utterances may be joined into a single *U* when certain pragmatic and phonological conditions are met; in addition, the *U*s must contract a certain syntactic and/or a positive semantic relation (for an exact definition of the phonological utterance and further details, see Nespor and Vogel 1986: Chapter 8).

As to our key questions in (4), the stress-sensitivity of lenition is present in the above analyses in the formulation of the rules, which make direct reference to the foot. Ordering aspiration before the two lenition rules guarantees that once a position is foot-initial and therefore strong, it always remains so. Since word-final consonants are never foot-initial, they will undergo one or the other lenition rule, depending on the cross-word situation. Word-initial unstressed syllables are generally thought to be adjoined to an adjacent foot, forming a superfoot with it, and thus word-initial plosives become foot-initial. As a consequence, both word-initial and prestress plosives are foot-initial. This is exemplified in (7) by the representation of the word *potato*, taken from Jensen (2000: 189).

(7)



As mentioned above, the type of lenition and its triggering environment are unconnected: simply, there are two different rules for flapping and glottalisation, manipulating rather ad hoc features like [release] or [tense], carefully ordered.

Function words exhibit special behaviour when they are unstressed because they are clitics: they cliticize onto their hosts either forming a clitic group (e.g. Nespor and

Vogel 1986) or getting incorporated into the prosodic word of their hosts (e.g. Jensen 2000). If they do not project feet themselves, their initial consonants will never be foot-initial and will be subject to lenition, as in *What do you want me to do?*.

The behaviour of /t/ in consonant clusters follows from its syllabification, much the same way as in the ambisyllabicity approaches described above in Section 3.2.1. As far as t's before syllabic consonants are concerned, Jensen (2000) himself admits (in footnote 18) that his analysis is unable to account for the differences, viz., glottalisation before syllabic /n/ (as in *button*) but flapping in all other cases (*bottom, bottle, butter*) cf. Section 3.1.3. above.[27]

3.2.3 Standard GP

GP's main motivation in analysing English t-allophony has been the search for a local cause as Harris also points out in the introduction to his chapter on consonant lenition (1994: 194). The identification of the 'weak' phonological positions, viz. the coda and the foot-internal intervocalic position, is a major achievement of phonological theory, but is still a mere observation unless a logical connection is found between these sites and the events they usually house. On the one hand, the segmental alterations observed are expected to naturally follow from the inherent strength or weakness of the position; on the other, the fact that the two types of weak position systematically trigger different changes needs to be accounted for. This latter point will receive more emphasis in CV/VC phonology.

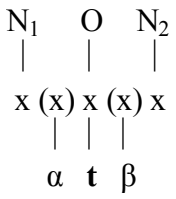
Harris and Kaye (1990) briefly survey the previous analyses, and conclude that all involve resyllabification, ambisyllabicity, or rule ordering – theoretical tools unavailable in GP (cf. Chapter 1). However, from these previous analyses they adopt the idea of breaking the lenition process into two stages, dating at least as far back as Kahn (1976) (see footnote 23 above). In conformity with GP's Element Theory and notion of lenition (Sections 1.5-1.6), tapping and glottalling are viewed as the loss of particular

[27] For Jensen's (2000) analysis of the "Withgott-effect", see Section 6.2.3.

elements from the internal composition of a stop. The less complex a segment is, the closer it is predicted to zero on some lenition trajectory. In sum, lenition is "progressive decomplexification" (Harris 1994: 123).

In Harris and Kaye's analysis, lenition takes place when a segment is attached to a position which intervenes within a governing domain. Since a governing nucleus possesses greater segmental licensing power than a licensed one, the former will support segments of a greater degree of complexity in its onset than the latter. The situation is sketched out in (8) (from Harris and Haye 1990: 258).

(8)



If we assume that N_1, being stressed, governs N_2, then the bold-faced /t/ in the onset position in (8) is unable to retain its complexity, therefore it will decompose, that is, lenite. Intervocalic and domain-final t̲'s are both sandwiched between two nuclear positions, followed by a licensed nucleus, as in (8); the difference lies in whether N_2 is phonetically realised or not.

Harris and Kaye aim to analyse both New York City (NYC) tapping and London glottalling, and observe that the sum of the contexts for the NYC tap and unreleased /t/ is equal to the sum of the contexts for the glottal stop in London. Consequently, the two phenomena must share a preliminary process, which they dub breaking: the elements composing the /t/ are rearranged into a contour structure. Breaking intrinsically feeds glottalling and tapping, as given in (9) (from Harris and Kaye ibid: 263): in London, the coronal element is deleted and glottalling ensues irrespective of what kind of nucleus

follows, whereas in NYC it is the stop element that disappears, yielding a tap, but only if the following nucleus is filled.[28]

(9)

Glottalling/tapping

```
Input   O       N
        |       |
        x       x
       / \
     ʔ°   R°
```

Output

	N empty	N filled
London: R° -> Ø	✓	✓
NYC: ʔ° -> Ø	*	✓

Cross-word tapping (as in *get on*) is derived cyclically: in the first cycle, all weakly licensed t̲'s undergo breaking, including word-final ones, as described above. Then in the next cycle the domain-final empty nucleus finds itself adjacent to the initial filled nucleus of the second word. At this point, the extended version of the OCP, alluded to in Section 1.1, becomes effective, and the empty nucleus is deleted. As a consequence, the broken /t/ is now followed by a filled nucleus, and accordingly tapping ensues. The

[28] Harris (1994: 214) also breaks tapping down into two stages, but in a slightly different manner. The first stage involves the deletion of the **h** element when the /t/ is followed by either an empty or a nonempty N (i.e., both in *bit* and *bitter*). The resulting object is interpretable as an unreleased stop. In the second stage further reduction takes place via the deletion of the **ʔ** element, yielding a tap, only if a nonempty N follows (i.e., in *bitter* only):

```
O   N          delete ʔ, iff x2
|   |          has melodic content
x1  x2
|
R
|
ʔ
```

That is, Harris banishes breaking and the contour structure of (9).

word-internal stress-sensitivity of tapping is expressed by the reference to the licensing potential of nuclei in the definition of breaking: only the onsets of prosodically recessive nuclei are affected, and no breaking occurs in pretonic position. Therefore no pretonic consonant serves as input to the rules in (9). Across words, however, by the time the final empty nucleus is erased to resolve the OCP violation, it has already induced the breaking of its onset /t/, and since stress relations are irrelevant to the rules in (9), all final t's are subject to them.

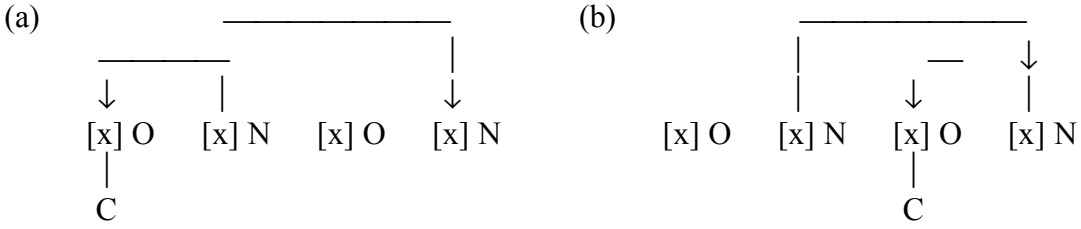
What unites word-initial and pretonic consonants is that neither are within a governing domain, that is, they escape breaking and thus lenition. In consonant clusters where the /t/ has governing obligations, it is shielded from lenition due to the Complexity Condition (Section 1.2). These are the so-called protected environments: the /t/ is only able to govern another consonant, and maintain the cluster, as long as it is at least of the same complexity as its governee, therefore it is protected from decomplexification, i.e. lenition. This applies to "coda"-onset sequences like the ones in *fist, left, fact, apt, after, custard, chapter, doctor* as well as branching onsets as in *petrol, mattress, patrimony*. It does not apply, however, in /rt/ clusters like *party* as /r/ is stipulated to occupy a nuclear position: the /t/ does not need to govern it since they do not form a consonant cluster. In *shelter* and *winter*, lenition is only possible when the /l/ or the /n/ is vocalized, i.e., when the /t/ is in fact intervocalic. The difference between words like *petrol* (with no lenition on the /t/) and *atlas, cutlass, chutney* (with glottalled /t/ in London) is that the latter contain fake (or bogus) clusters: the members do not contract a governing relationship but are instead separated by an empty nucleus. That is, *at* and *atlas* or *cut* and *cutlass* are expected to exhibit parallel behaviour.[29]

Harris's (1992a, 1994) theory of Licensing Inheritance is an improvement upon the Harris-Kaye analysis in several respects. It is based on the assumption that a licensed position inherits its a(utosegmental)-licensing potential from its licensor (Harris 1992a: 22, 1994: 206). Therefore, the length of licensing paths is of high significance: the longer the path, the more diminished the a-licensing potential, as illustrated in (10)

[29] For the discussion of two successive potential lenition sites in words like *competitive*, mentioned in Section 3.1.3 above (Harris and Kaye 1990: 261), see Section 6.2.X.

(from Harris 1992a: 33). While a foot-initial onset is at one remove from the ultimate licensor of the stress domain (10a), a foot-internal onset is at two removes (10b).

(10)



Word-internally, then, the stressed vowel is the head of the stress domain, which is unlicensed itself, and therefore has the strongest a-licensing potential, transmitted to its onset (10a). Unstressed nuclei, however, get their licensing from the stressed one and consequently they and their licensees will always be at one more remove from the domain head (10b). It follows that foot-initial onsets can support more complex segments than foot-medial ones.

A word-final consonant is at two removes from the ultimate source of licensing, in the same way as its foot-internal peer, as it is licensed by a final empty nucleus, which is itself licensed by parameter. Across words, Harris repeats Harris and Kaye's cyclic analysis involving the suppression of the final empty nucleus (to satisfy the extended OCP)[30]. It is worth mentioning that in such an analysis phrases like *get by* are derived completely differently from phrases like *get a*: the final empty nucleus remains in the structure in the first case, but it gets deleted in the second case.

Harris (1994: 224-225) also treats /t/ before syllabic resonants, more specifically, before syllabic /n/. He claims that in *button* and the like the vocalic content of the unstressed nucleus is usurped by the nasal spreading into its position, and when the syllabic consonant is created, the OCP requires the coalescence of the syllabic consonant and the preceding /t/, since both contain the stop element and the coronal

[30] Harris (1994: 204) defends the deletion analysis by relaxing Structure Preservation so as to refer to "general conditions on licensing" which are required to remain in force throughout derivation. Here, the relevant general condition is Onset Licensing, and the final consonant in *get* is indeed licensed by a following nucleus in both *get* and *get a*: by the final empty nucleus in the first case, and by the filled nucleus of the article in the latter.

element. Sharing melody with a following position protects the /t/ from decomplexification: no tapping or spirantization is expected. Unfortunately Harris is silent about how it is nevertheless possible to lenite the /t/ before other syllabic consonants. Recall from Section 3.1.3 above that it is syllabic /n/ only that systematically resists the tapping of a preceding /t/, but tapping is possible with syllabic /l/, /m/ and /r/ (*bottle, bottom, butter*). In Harris' system, /l/ also contains both **?** and **R**, so it is predicted to behave in the same way as /n/, contrary to fact. /m/ and /r/ also share one element each with /t/, namely **?** and **R**, respectively – perhaps one shared element is not enough to protect the /t/ in this case, although the number of primes shared does not usually affect process inhibition, as is evident from, e.g., Honeybone (2005).

As pointed out in Szigetvári (1999b: 45-46), to unite the licensing paths of the two types of coda (viz., word-internal and final), Licensing Inheritance must stipulate that medial "codas" (i.e., post-nuclear rhymal adjuncts) are only licensed by the following onset (in accord with the "Coda" Licensing Principle). Crucially, they are not licensed by their nuclei within the rhyme they share, because in that case posttonic codas (at one remove from the metrical head) would be stronger than the others (two or more removes). However, if we assume that word-internal codas only receive licensing from following onsets, they too are at least two removes from the ultimate licensor. But then the only syllabic constituent in which the head does not license the dependent is the rhyme – in contrast to the onset and the nucleus.

From the Licensing Inheritance analysis it is not clear what word-initial and pre-stress positions have in common that makes them both strong. In a word like *potáto*, if the first nucleus gets its licensing from the stressed vowel in the same way as the word-final one, then the initial /p/ is at two removes from the domain head in the same way as the second, foot-internal /t/, and therefore both are expected to undergo tapping. Szigetvári (1999b: 46-47) devotes some space to pondering about how Licensing Inheritance could be accommodated to cover word-initial nonlenition. He concludes that promoting the initial syllables of words like *potato* and *today* to foot status solves the problem: their vowels become the ultimate licensors, so no difference is expected

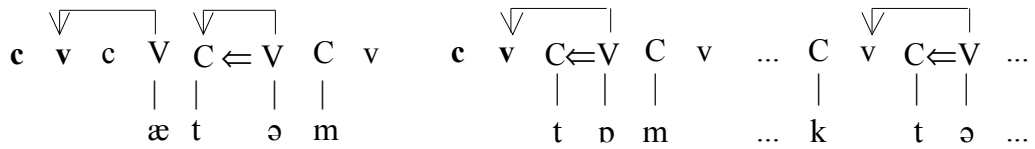
between their onsets and pretonic consonants. The problem with such an analysis is that allowing the vowels of footheads to reduce to schwa (which they do indeed in the examples under scrutiny) seems to be a hasty relaxation of the theory, and it remains unclear why the vowels of other monosyllabic feet are unable to do the same (e.g., in the final syllable of words like *hesitate*). One must state that the reduction of foothead vowels is only possible word-initially, that is, special reference to the left edge of the word is needed anyway.

As a final evaluation, we find that neither of the Standard GP analyses explain why there are different types of lenition in foot-internal intervocalic positions and in codas, how function words differ from lexical content words, and what the domain of flapping is.

3.2.4 CV/VC phonology

Strict CV/VC phonology, conceiving of prosodic structure as a sequence of strictly alternating C and V positions, bases its theory of lenition on the lateral relationships skeletal slots contract. In Section 2.1.2, Ségéral and Scheer's (1999a) redefinition of government and licensing in the theory of Coda Mirror (CM) was introduced. The two are antagonistic forces emanating from filled V positions: Proper Government inhibits, but licensing comforts, the segmental expression of its target (Ségéral and Scheer 1999a: 20). Also, recall from Section 2.1.1 Lowenstamm's boundary-marker, the empty cv-span marking the left edge of words. These theoretical tools add up to yield the representations in (11), repeated from Section 2.1.2 for ease of reference.

(11) a. *atom* b. *Tom* c. *Acton*



As (11) shows, the intervocalic /t/ is expected to lenite, as it is governed by the following vowel (11a), whereas a word-initial /t/ occupies a strong phonological position, as it is licensed only (11b). What is common to the two subcases of Coda Mirror, viz. word-initial (11b) and postconsonantal positions (11c), is the strength of the /t/ due to the fact that the following vowel has to govern an empty v to the left of the /t/, and thus the /t/ itself escapes government.

Szigetvári (1999b: 53-55) compares the predictions that Licensing Inheritance and CM make with respect to consonant lenition[31], and finds that while Licensing Inheritance suffers from the inability of distinguishing between the two types of lenition, CM is also incomplete since it has no account of prestress strength, i.e., of the stress-sensitivity of lenition (at least in certain languages)[32]. He identifies a site where the two models diverge, **V**C_V[33], where Licensing Inheritance expects lenition (at two removes from the boldfaced licensor), CM, however, correctly predicts its absence (in a sequence ...**V**Cv_1<u>C</u>V_2..., unstressed V_2 must properly govern v_1 to silence it, therefore the underlined C is licensed only). An essential difference between the two theories, Szigetvári observes, is the conception of licensing: for Licensing Inheritance it is a scalar property (gradually decreasing as the licensing path grows), for CM it is binary. Although Szigetvári evaluates this as a weakness of Licensing Inheritance, in Chapter 6 we will find that a distinction in the degree of weakness of weak phonological positions is indeed called for, but it can be expressed in the spirit of CM.

[31] Licensing Inheritance is also evaluated in Scheer (2004: 218-230). One of the major conclusions is that although stress is indeed a conditioning factor in consonant lenition, it is a secondary factor, since it only plays a role in a subset of lenition systems. Therefore, it is undesirable to build it directly into the derivational mechanism, as is done in Licensing Inheritance. Instead, it should be present in the grammar in the form of a parameter. This is what we find in CM+ and the present thesis (see Chapter 5).
[32] Also admitted by Scheer (2004: 118).
[33] The boldfaced V is stressed, the other is not.

Ségéral and Scheer (1999c) (cited in Szigetvári 1999b: fn.100 p.95) argue that stress gets materialized in the form of an empty cv unit, triggering Italian Tonic Lengthening and aspiration in English: it takes place when the plosive is preceded by an empty cv span.

A development upon CM was described in Section 2.1.3: Coda Mirror Plus (CM+), combined with VC Phonology (Dienes and Szigetvári 1999, Szigetvári 1999b and Dienes 2000), is an elaboration on CM, including concrete definitions of the inherent properties of the skeletal slots ("vocalicness is loud, not only acoustically but also in the sense that V slots in the phonological skeleton aim at being pronounced" while "consonantalness is mute, if nothing intervenes a C position will stay silent") (Szigetvári 1999b: 62) along with a proper circumscription of the effects the two lateral relations impose upon them. Ségéral and Scheer's definition of licensing is now accompanied by that of government, which is assumed to spoil the inherent properties of its target (Szigetvári 1999b: 66).

Section 2.2.1 above spent some space on the difference between a skeleton consisting of VC units and one comprising CV spans. In the analysis of lenition, VC units come in handy in the expression of the stress-sensitivity of the process. Dienes and Szigetvári observe that stressed vowels are unable to govern, either their "onsets" or another nucleus to the left. To them, this means that a stressed V's government cannot penetrate into another VC unit. Since in English feet are trochaic, this in turn means that government cannot penetrate a stress domain. This is the so-called Antipenetration Constraint (e.g., Szigetvári 1999b: 79).[34] It follows that while foot-internally the /t/ is both governed and licensed (i.e., it is able to support a relatively complex segmental structure but its inherent muteness is spoiled forcing it to sonorize, which predicts tapping in that environment) (12a), and all initial t̲'s are strong because they evade government (12c), pretonic consonants are also strong (at least in languages like English) because they reside outside the following vowel's stress domain (12b).

[34] The Antipenetration Constraint enriches the theory by being parametrically present or absent in languages, thus differentiating between stress-sensitive and stress-insensitive lenition systems. (Szigetvári ibid: 137)

(12) a. (*atom*) b. (*at*)(*omic*) c. (*vTom*)

```
V  C⇐V  C        V  C⇐V  C  V  C        v  C⇐V  C
|  |   |  |        |  |   |  |  |  |        |   |  |
æ  t   ə  m        ə  t   ɒ  m  ɪ  k        t   ɒ  m
```

Word-internally, then, a difference is predicted between stressed and unstressed vowels: since the former start a stress domain, they are stopped from emitting government by the Antipenetration Constraint and the pre-stress consonant stands in a strong position. As is alluded to in Section 2.2.2 and elaborated on in Chapter 5, the VC skeleton predicts the same difference across words, which is not borne out by the data. Unfortunately, neither Dienes nor Szigetvári provides the treatment of tapping at the phrasal level. For further discussion, see Chapter 5.

The fact that there are different types of lenition in foot-internal intervocalic positions and in codas is straightforwardly accounted for in CM+, which distinguishes two types of phonologically weak positions. The first is before an empty v or nothing[35], which is roughly before a consonant and word-finally (that is, the classical "coda"); in such cases consonants remain ungoverned and unlicensed, therefore they are able to keep up their inherent muteness but are prone to decomplexify, i.e., they exhibit consonantal lenition (t̲'s are glottalised). The second type of weak position is the one occupied by foot-internal intervocalic consonants, which receive both government and licensing from the following (unstressed) vowel (as in (12a)); here consonants tend to move towards vocalicness, e.g. GA t̲'s are flapped.

CM+ and VC phonology distinguish three types of consonant cluster, essentially on the basis of their distribution: (i) so-called onset clusters, which appear initially and medially in many languages but are marked finally; (ii) so-called coda clusters, which appear medially and finally in many languages but are marked initially; (iii) so-called bogus clusters, which appear only medially in most languages including English. The first group typically contains rising-sonority clusters like /tr/, the second includes

[35] At this point, CV and VC phonology do not diverge significantly. In CV phonology, all "codas" are followed by an empty nucleus, which is unable to license or govern medially and parametrically set to either license or govern (or both) finally. In VC Phonology, only medial "codas" are followed by an empty v, word-final C's are unlicensed and ungoverned because they are followed by nothing.

sequences with falling sonority such as /st/ and coronal-final plateaus like /kt/, while the third is the "wastebin" of the rest, e.g. /tl/, /tk/, /tn/, /gm/, etc. The members of onset clusters contract a right-to-left licensing relation, which silences the intervening empty v on the one hand (which is inert as a consequence), and creates a strong position for the first consonant on the other (cf. (13a)). The same holds true of the second C in coda or bogus clusters, which occupies the licensed position of the configuration shown in (11c) above. All of these consonants are expected to be aspirated if they are voiceless plosives, as in *trip, actor, Atkins*.

(13) Possible consonantal positions (Szigetvári 1999b: 135, chart (95)[36])

	LIC'D	GOV'D	LENITION TYPE	POSITION
a.	yes	no	none	#_, oc1, bc2, cc2, _**V**
b.	no	no	c-lenition	_#, bc1
c.	yes	yes	v-lenition	V_*V*
d.	no	yes	c-/v-lenition	cc1, within a long V

As the consonants participating in coda clusters enter into a right-to-left governing relationship, the first member finds itself in a very special situation: it is governed (by the second member) but not licensed (because it is followed by an empty v). That is, such consonants are expected to undergo either sonorization (since they are governed) or melodic loss (since they are unlicensed) or a combination of the two. This is borne out by the data: the first member of coda clusters is a sonorant ideally, which also very often sacrifices its place element in favour of that of the following consonant. (Szigetvári ibid: 135) Unfortunately CM+ has no story of how aspiration is bled in *sC* and *sCC* clusters.[37]

Besides the issue of cross-word tapping, CM+ does not address that of how function words differ from lexical content words, what the domain of flapping is, and how the behaviour of /t/ before syllabic consonants is explained. It does suggest that syllabic consonants are most probably the interpretations of VC sequences where the

[36] Abbreviations: bc*n*, cc*n* and oc*n* mean the *n*th position in a bogus, coda and onset cluster, respectively; **V** is a stressed, *V* is an unstressed vowel.
[37] For more on consonant clusters in CV/VC phonology, see Sections 5.3.4-5.3.5.

two positions are simultaneously associated with the consonantal melody, and the V position is able to govern and license (Szigetvári ibid: 117, 159ff.), however, this representation is not related to t̲-lenition in any way[38]. Notice that it predicts tapping before all syllabic consonants: in C_1VC_2, where VC_2 constitutes the syllabic sonorant, the V both governs and licenses C_1, therefore vocalic lenition is expected. Thus tapping in *bottle*-type words, but not glottalisation in *button*, is accounted for.[39]

3.2.5 Conclusion

The above sections surveying some of the previous analyses of t̲-allophony show their partially overlapping coverage of the empirical observations: there exists a common core of phenomena that is subsumed by all (with more or less explanatory adequacy), but the "periphery" contains a whole lot of issues with respect to which one or the other theory fares better, or which is explained by none. What Chapter 5 attempts to achieve is putting the strengths of these accounts together like the pieces of a puzzle, making the necessary modifications, and by doing so, arriving at a description which addresses all the (systematically observable) facts enumerated in Section 3.1 above.

[38] Scheer (2004: 283ff, esp. 309ff) argues for the same representation, i.e., that syllabic consonants are left-branching VC structures, and he too draws no distinction between their types, and does not examine their role in lenition. Surprisingly enough, after some discussion of the behaviour of syllabic sonorants, Scheer (2004: 362) tentatively proposes that they are in fact both left- and right-branching, i.e., they occupy a VCV string.

[39] In fact, Szigetvári (1999b: 159ff.) argues that /r w j/, the glides of English, are syllabic differently than other consonants; this would derive a distinction in behaviour between syllabic /r/ (as in *butter*) and the rest.

Chapter 4: The boundary-marker

The fact that the left edge of (phonological) words is a strong position counts as a phonological commonplace. This basically means that the beginning of the word favours fortition processes and disfavours lenition both synchronically and diachronically. Theories have usually attempted to account for this with reference to the word boundary (#) or to foot-initial position. Recall from Section 2.1.1 that, as an alternative, most practitioners of Strict CV Phonology assume that each word of a major category begins with a melodically empty CV unit on the skeletal tier, marking the word boundary (after Lowenstamm 1999). One of the functions of the boundary-marker (emboldened) in a **cv**CV... word (that is, a word starting with a single consonant followed by a vowel, where lower-case letters denote empty skeletal positions) is to absorb the reductive force of government emanating from the first vowel of the word, thus the empty v in the boundary-marker will be prevented from being pronounced, and the word-initial phonetically expressed consonant will not be negatively affected, i.e., it will not lenite (see below).

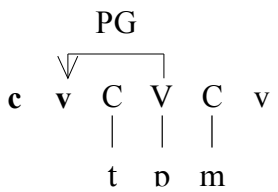
So far, the study of this boundary-marker has concentrated on the behaviour of consonant-initial words, therefore this chapter has two main aims. On the one hand, it investigates whether or not vowel-initial words also possess a boundary-marker, and what the difference is, if there is one, between consonant- and vowel-initial words; on the other hand, it looks into what happens to the boundary-marker post-lexically, i.e., in connected speech. It is structured in the following way. Section 4.1 introduces the basic language typology predicted by the presence vs. absence of the left-edge boundary-marker. Section 4.2 summarises the insights of Prosodic Phonology (e.g., Nespor & Vogel 1986), and in Section 4.3 it is suggested and exemplified that the boundary-marker serves as a general boundary marking the edges of (all) phonological domains: certain phonological rules will arbitrarily decide to ignore it and treat it as a kind of extraprosodic skeletal material. The boundary-markers not ignored by a given rule will delimit its domain by blocking its application. Then, in Section 4.4, a typology of the

effects of the extraprosodic (i.e., phonologically inert) boundary-marker is provided, which highlights the special status of the situation when a consonant-final word meets a vowel-initial one. Several examples are given, mostly from English. Section 4.5 discusses language typology and parameters, and finally, in Section 4.6, further issues are addressed.

4.1 Preliminaries

Recall that the boundary-marker has at least two functions. First, it explains why the beginning of the word is a strong phonological position. Check in (1) that the /t/ in *Tom* escapes government because the ECP requires PG to hit the empty v in the boundary-marker.

(1)

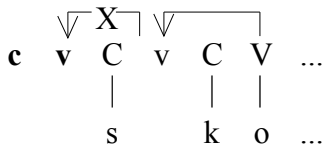


Second, the set of word-initial clusters accepted as licit in a language can also be described with reference to the boundary-marker. In certain languages, e.g. Latin, certain falling-sonority consonant clusters regularly occurring word-internally are also attested in the initial position: /s/ plus obstruent clusters are notorious for that in other Indo-European languages too. In the present theoretical framework we have to say – though the reason is not well understood – that the ECP is exceptionally suspended for the boundary-marker in those cases, or, as Kaye (1992) puts it in Standard GP, the empty v preceding /s/ is "magically" licensed (cf. (2a)). During their development, however, several Romance languages (Spanish, French) "reimposed" the general interpretation of the ECP, and the ungoverned empty v in the boundary-marker had to surface. Therefore,

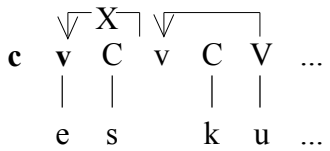
Lat. *schola* > Sp. *escuela* / Fr. *école* (with further changes – see, e.g. Ségéral and Scheer 1999b) (2b).

(2)

a.

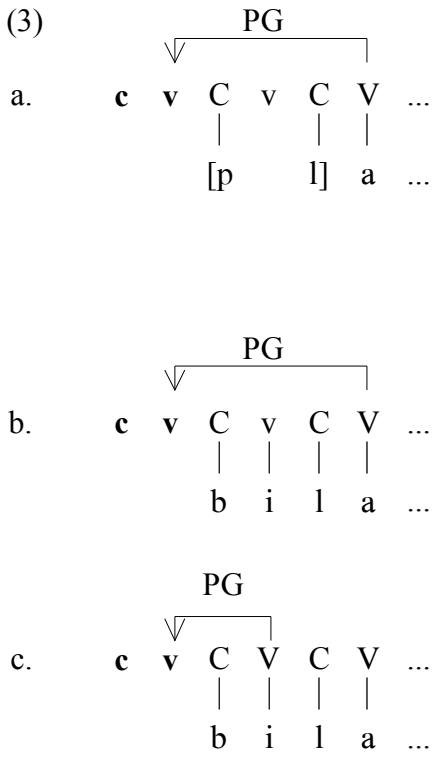


b.



Besides prothesis, anaptyxis is another repair strategy used in situations when forms are adopted from one historical state to another, or one language to another, across dissimilar phonotactic systems. (3) illustrates loanword adaptation in Egyptian Arabic (EA): *bilastic* 'plastic' has the interconsonantal empty vowel, silenced in English by virtue of being contained in a closed domain (cf. Scheer 1996, Lowenstamm 1999), pronounced in EA. It is a mystery why EA (and a lot of languages – cf., e.g., Fleischhacker 2001) chooses anaptyxis rather than prothesis in word-initial rising-sonority clusters (abbreviated by some to #TR); one possible explanation is that somehow the governing relations in the source language are also borrowed along with the form, as shown in (3b), and only after the word has become integrated into the system do those relations get restructured (3c).

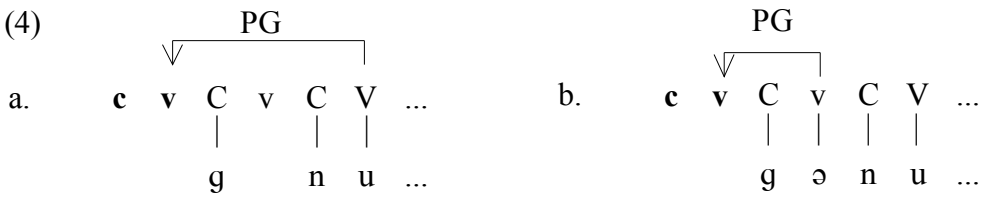
(3)



In English, the most frequent strategy is the deletion of the first consonant in word-initial coda-onset sequences other than *sC*, cf. *gnu, psycho, pneumonia, pterodactyl, mnemonics, Xerxes, xylophone*, etc. The boundary-marker attracts the PG emanating from the pronounced vowel, and since these clusters are unable to form a closed domain, the sandwiched empty v remains ungoverned, unable to license the /g/, which deletes as a consequence (4a).[1] Apparently, an ungoverned empty v position which does not receive phonetic interpretation for some reason has a reduced licensing capacity compared to its governed counterpart: in word-internal position, where the empty v between /g/ and /n/ *is* reached by the PG of the following vowel, it is "strong" enough to support the consonant to the left as in *magnet*.

[1] Notice that it is still puzzling how English *sC* clusters survive simplification.

(4)



(4b) illustrates the anaptyxis solution resulting in an alternative pronunciation of *gnu* (which is judged to be "jocular" in LPD).

Even in the situation in (1) the empty v of the boundary-marker may manifest itself. In many languages there are constraints on what (single) consonants can appear word-initially, e.g., liquids are frequently restricted to intervocalic position. In the case of loanwords from languages not having the same restrictions it very often happens that a prothetic vowel appears at the beginning of the loanword to avoid initial liquids. Examples and a representation are given in (5).

(5) Word-initial prothesis before a single consonant[2]

a. Sanskrit *lok*- 'world' > Tamil *ulakam*

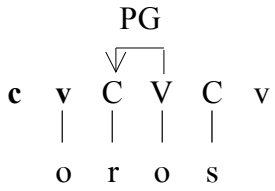
Sanskrit *rAj*- 'king, prince' > Tamil *aracan*, *arayan*

Latin *reg*- 'king' > Basque *errege*

rus-, *ros*- 'Russian' > Altaic *urus*-, *oros*- (> Hungarian *orosz*)

> Old Japanese *orosia* 'Russia'

b.



[2] I am grateful to Ivan A. Derzhanski and Laurence Labrune for providing me with examples.

In (5b), the liquid requires to be governed, i.e., to occupy intervocalic position. Consequently, the full vowel is unable to govern the boundary-marker, whose empty vowel receives phonetic interpretation.

According to the workings of the boundary-marker just explained, two basic language types are predicted. On the one hand, modern occidental Afro-Asiatic (Algerian, Tunesian, Moroccan Arabic, Berber), Slavic and Greek have been shown to allow for any combination of consonants as well as lenition word-initially (see, e.g., Scheer 2001 for Slavic[3], and Seigneur-Froli 2004, to appear for Greek). Such languages will henceforth be referred to as "permissive" (Scheer 2001 uses the name "anything-goes" languages, which is even more expressive; Seigneur-Froli (to appear) uses the adjective "free"). Other languages like French or English, however, display a strong preference for TR-clusters word-initially, and no lenition is expected at that location. We can call these languages "strict" (Scheer's "#TR-only" languages, Seigneur-Froli's "restricted" category).[4] [5]

The proper way to distinguish these two language types is one of the main concerns of this chapter. So far, two options have been pursued in the literature. Lowenstamm (1999) represents the original stance claiming that the boundary-marker is *always* licensed (i.e., it always requires PG) in strict languages like Spanish (that is why certain clusters such as the one in (2a) are prohibited word-initially) whereas it is *not always* licensed in 'permissive' languages like Biblical Hebrew (depending on the cluster in question) (6).

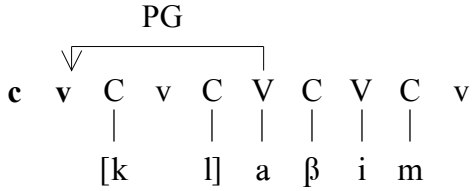
[3] More specifically, Czech, Polish, Russian, Ukrainian, and Serbo-Croatian are classified as permissive in Scheer (to appear), whereas Bulgarian, Macedonian, Slovak, Slovenian, and Bielorussian are claimed to be strict. As we will see in Section 4.5 below, Bielorussian exhibits exceptional behaviour with respect to the distribution of the boundary-marker.

[4] As the discussion unfolds, we will find that being a strict language and keeping the beginning of the word unaffected by weakening, at least from a synchronic perspective, do not go hand in hand. Certain strict languages *do* allow for the lenition of word-initial consonants across word boundaries, e.g. Italian.

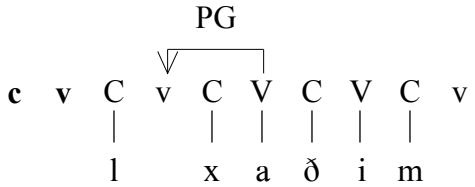
[5] As Scheer (to appear) remarks, this binary classification of languages into strict and permissive is to highlight a phonologically relevant distinction rather than hide or ignore the cross-linguistic continuum initial cluster types exhibit. In addition, it reflects the set of *potential* initial clusters in a given language, and not the actually attested inventory: as Scheer (2001) demonstrates for Slavic, nonoccurring #RT-clusters are accidental gaps and judged acceptable by native speakers.

(6) Biblical Hebrew (data from Lowenstamm 1999: 159-160)

a. klaβim 'dogs': the initial cv is licensed

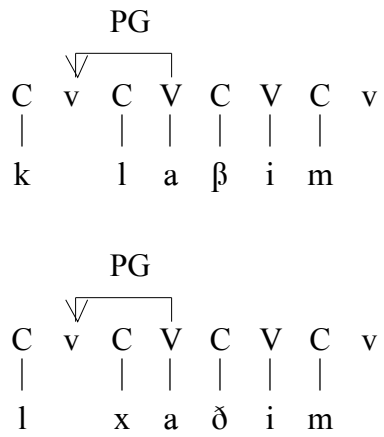


b. lxaðim 'captures (n)': the initial cv is not licensed



In contrast, according to Scheer (2001), the distinction lies in the presence vs. absence of the boundary-marker: in strict languages it is present and needs licensing while in permissive languages it is absent. The difference between Lowenstamm's and Scheer's conceptions of the boundary-marker is apparent in how they analyse word-initial clusters in permissive languages. For Scheer, all such clusters are "fake" or "bogus", irrespective of the consonantal melodies involved, i.e., they never form a closed domain. Consequently, the sandwiched empty v always requires government, which would never reach the boundary-marker should it be there. If we try to adapt Lowenstamm's Hebrew examples to Scheer's model, the representations in (7) ensue.

(7)



As Scheer (2004, to appear) and Seigneur-Froli (to appear) point out, Lowenstamm's version violates the ECP since it leaves certain empty v's ungoverned, i.e., unsilenced. Scheer argues on theoretical grounds: morphological information in phonology should be privative – linguistic objects are either present or absent, and accordingly, the beginning of the word is either projected into phonology or it is not. Thus the well-known absence of languages having bogus initial clusters only, to the exclusion of #TR clusters, is also predicted, although in Lowenstamm's framework it is a theoretical possibility to have words like (6b) only, in a language in which the boundary-marker is *never* licensed. In addition, in such a language no words would start with a single consonant, since in the configuration in (1) the boundary-marker *is* licensed, thus such a case would never yield in a language where the boundary-marker is *never* licensed.[6] Lowenstamm himself considers this option, and concludes that the observed facts suggest that the correct statement of the parameter makes reference to licensing *always* or *not always* taking place. Clearly, this is a theoretical question, closely connected to our understanding of the nature of grammar.[7]

[6] Notice that according to Lowenstamm's logic, cv's are always licensed whenever they *can* be licensed, i.e., a governor is available; and they are only left unlicensed (by parameter) when no such licensor is available. Consequently, since a governor is always available in the case of single initial consonants, the boundary-marker is always licensed there.

[7] At one point Scheer (to appear) himself admits that this may be a mere matter of vocabulary: there is no empirical evidence either for or against the existence of the boundary-marker in permissive languages. Although Lowenstamm (1999) refers to cross-word gemination and lengthening in Biblical Hebrew as an

This chapter concentrates on the analysis of strict languages, but as far as language typology goes, it essentially takes Scheer's stance as a starting point and modifies it later, only to finally fuse it with Lowenstamm's into a kind of theoretical compromise (in Section 4.6). But first, let us examine some other aspects of cross-word phenomena, with examples from strict languages.

4.2 Prosodic domains

The theory of the domains of phonological processes in Generative Phonology was launched in SPE by introducing two boundaries, the formative boundary (symbolised by +) and the word boundary (#), presumed to have relevance in all languages. In what follows, we will concentrate on the latter.

The general convention governing the appearance of the word boundary in the phonological surface structure is given in (8):

(8) The boundary # is automatically inserted at the beginning and end of every string dominated by a major category, i.e., by one of the lexical categories "noun", "verb", "adjective", or by a category such as "sentence", "noun phrase", "verb phrase", which dominates a lexical category. (SPE: 366)

The above mechanism inserting #-boundaries applies at the syntactic surface structure (represented with labelled bracketing indicating categorization) and generates a representation which is then modified by readjustment rules replacing some occurrences of # by + as well as deleting (and perhaps also introducing) some in various positions.

---

argument for an empty cv between words, providing the landing site for the spreading melody, Scheer denies that Biblical Hebrew is a permissive language at all (also in Seigneur-Froli to appear). As far as the status of Biblical Hebrew is concerned, I would not like to take a stance here, although the arguments are rather convincing: in words with initial RT-clusters like *lxadm* 'captures (n)' (see (6b) above) the classical writing exhibits a schwa between the consonants (i.e., *l@xaðm*), while this is not done word-internally. Despite the fact that the exact phonetic realisation of this schwa is subject to debate, its presence is evidenced by the (regularly postvocalic) spirantisation of the following consonant, which results in the /x/ of the above example word from the /k/ of the root *lkd*.

The resulting phonological surface structure is the one that enters the phonological component of the grammar.

The application of the convention in (8) also provides a definition of the notion "word", which can be roughly given as a string of formatives sandwiched between pairs of #-boundaries (i.e. ##_##). (For a more precise definition see SPE: 367.) Ignoring the exact number of boundary symbols, later # became shorthand for word edges, featuring in the famous disjunctions _{#, C} and {#, C}_, the frequent contexts for phonological processes. However, the reintroduction of hierarchical structure (i.e. the syllable) into supraskeletal representation (e.g. Kahn 1976 for English, identifying {#, C} as a notational artefact standing for the syllable boundary) and its advent in subskeletal melody (Autosegmental Phonology – Goldsmith 1976) forced analysts to reinterpret phonological domains.

Among other aspects of the linear framework of phonology, word boundary theory was found inadequate as laid down in SPE. (For a discussion of its disadvantages, see Kaisse 1985: Chapter 5). A new branch of phonological theory, Prosodic Phonology (henceforth PP), was introduced to account for the syntax-phonology interface, the way the syntactic and phonological components of the grammar are organized with respect to each other. According to PP, "the mental representation of speech is divided into hierarchically arranged chunks" (Nespor and Vogel 1986: 1), the so-called prosodic constituents (see (9)). The basic idea is that syntax does not provide domains for phonological rules in a direct fashion, but another level of representation, dubbed "p-structure", must be posited (for the arguments, see e.g. Nespor and Vogel 1986). P-structure mediates between syntactic surface structure (frequently referred to as s-structure) and the phonological module, and functions as the locus of their interaction.

The difference between the SPE-type early version of p-structure and the one proposed by PP lies in the fact that while in the former p-structure (more or less) directly derives from s-structure, and the only syntactic property governing the syntax-phonology mapping is constituency and phrasal rank expressed by the number of boundary symbols, in the latter the two representations are distinct, not (necessarily)

isomorphic, and only indirectly related. Researchers do not agree on the nature of the syntax-phonology mapping; for a summary of the different mapping algorhythms proposed see Inkelas and Zec (1995).

Prosodic Phonology (PP), then, claims that there exists a hierarchy of prosodic constituents which serves as the inventory from which the rules choose their domains of application. Although authors slightly differ as to what these constituents are, the common core of all models is given in (9).[8]

(9) (mora)
(segment)
(syllable)
(foot)
phonological word
(clitic group)
phonological phrase
intonational phrase
utterance

PP is therefore a theory of phonological domains, a subsystem of the phonological component of the grammar that organises strings of language into these phonological units, which constitute the domains of application of phonological rules.

The most convincing piece of evidence for the inevitability of PP comes from cases when the application of a given phonological rule depends on nonphonological (mainly syntactic) information: under identical segmental conditions, for example, French liaison applies in phrases of a certain type (as in Booij's (1986) examples in (10a) and Nespor and Vogel's (1986: 41) examples in (11a)) but fails to do so in phrases of a different type (as in (10b) and (11b)), where final consonant deletion applies instead.

[8] The phonological units given in parentheses feature only in some of the analyses written within the framework of PP; some of them, e.g. the clitic group, are even rejected as a prosodic constituent by most authors.

(10) a. un [savant]$_A$ [anglais]$_N$ 'a learned Englishman'

b. un [savant]$_N$ [anglais]$_A$ 'an English scientist'

(11) a. Les giraffes et les éléphants sont ses [meilleurs]$_A$ [amis]$_N$

'Giraffes and elephants are his best friends'

b. Claude a des [perrochets]$_N$ [intolérables]$_A$

'Claude has some intolerable parrots'

It is also evident that rules select their domains of application arbitrarily. Rules with similar structural descriptions and changes may apply within different domains, as is the case of final consonant liaison in French (within the phonological phrase) as opposed to r-liaison in English (within the utterance). Even the same phonological rule may choose different domains in different dialects of the same language: in English, l-darkening applies within the utterance in RP whereas it applies within the word in several American dialects.

In CV phonology, the left word boundary is marked by the empty CV unit. If it is the boundary-marker that makes the beginning of the word a strong phonological position, it means it blocks the application of lenition rules (where "rule" of course means something like the interplay of forces like government and licensing – a "declarative" rather than a "procedural" sense of the word). It follows, then, that this empty skeletal unit can be concieved of as a general boundary-marker which circumscribes a given rule's domain of application, at least in the case of segmental alternations[9], and rules taking constituents larger than the foot as their domain.

Nespor and Vogel (1986), following Selkirk (1980), distinguish three rule types: domain span, domain juncture and domain limit rules. *Domain span rules* apply within the boundaries of a particular prosodic constituent in such a way that the triggering environment is contained somewhere within that constituent. An example of this rule type is flapping in English (cf. Chapter 3), which is a U-level rule, i.e., it applies within

[9] Suprasegmental features such as stress or tone are not considered here.

the phonological utterance. One of its most frequent triggering environments is intervocalic position, i.e., V_1__V_2, and it only applies when both V_1 and V_2 are found within the same utterance. Such rules are analysed here in a rather simple manner: any occurrence of a boundary-marker representing an utterance boundary either to the right of V_1 or the left of V_2 will block the application of the rule. The prediction made by CV phonology is that most cases of domain span rules will be segmental lenition processes caused by the government emanating from a following nonempty vocalic position. Whenever a boundary-marker intervenes between this vowel and the target, the pressure to silence the v part of the boundary-marker via PG will be stronger and no government charge will be left to hit the target (cf. the detailed analysis of English flapping in the following chapter).

Domain limit rules are quite similar to domain span rules; the difference is that in this case the triggering environment must be situated at the *edge* of a given prosodic constituent. Mostly these are either processes affecting segments syllable-initially (i.e., before a nonempty vowel) and foot-initially (stress-sensitive fortitions like aspiration in English – see Chapters 3 and 5), or lenitions in final positions. In English, t-glottalisation is analysed in PP as a domain-limit rule applying syllable-finally (Nespor and Vogel 1986: 78); RP l-darkening (velarization of /l/ before a consonant or a pause) and the r-dropping rule of non-rhotic dialects, on the other hand, are utterance-limit processes (although Nespor and Vogel (ibid: 226ff) do not recognize r-dropping as a separate rule but integrate the phenomenon into a general rule of r-insertion taking place within the utterance). Other examples include p-word-final lenitions like Final Cluster Reduction, Final Deaspiration and Devoicing in Sanskrit (Selkirk 1980, Nespor and Vogel 1986: 118). Sanskrit also exemplifies weakening in absolute final position: Visarga at Pause (Selkirk 1980, Nespor and Vogel 1986: 230) lenites /s/ and /r/ to become /h/ utterance-finally. In all these cases, the limited licensing abilities of FENs manifest themselves: they are unable to support r's, coronal l's and t's in English, clusters, laryngeal features, and, again, coronals in Sanskrit. As we will see in the rest of this chapter as well as in the following one, across words languages have the choice as to when a following word-initial segment is permitted/enabled to affect a preceding

consonant. In the case of utterance-bounded rules, it always is within the whole utterance; in the other cases certain boundary-markers act as obstacles to that influence.

The third rule type recognized in PP, *domain juncture rules* are, as the name suggests, processes triggered by the presence of the domain juncture, i.e., the boundary-marker itself. Here CV phonology expects to primarily find segment fortition and lengthening rules. Fortition, being the mirror image of lenition, is predicted to occur next to a boundary-marker exactly because of its blocking effect hindering government, as described above. In addition, consisting of empty skeletal slots, the boundary-marker can also serve as the landing site of some melody spreading from either direction – that is why lengthening is expected[10]. Examples are provided by Raddoppiamento Sintattico, which is a p-word juncture rule on the Phonological Phrase domain (i.e., a rule triggered across the p-word boundary within the limits of the Phonological Phrase; cf. Nespor and Vogel 1986: 165ff)[11], or Somali gemination, where stem-final velars spread onto the empty skeletal position of the causative suffix, and therefore undergo palatalization (see Scheer to appear). We will see below, however, that very often there are additional conditions on the spreading of some melodic material, so the mere presence of the boundary-marker may not always be sufficient to trigger the process (as in the case of French liaison, discussed in the following section).

Unfortunately, there are a number of processes (other than lengthening rules) which apply across word boundaries but not within words. A number of assimilations are like that, e.g., the Sanskrit rule of Stop to Nasal (Nespor and Vogel ibid: 118) or Greek Nasal Assimilation and Stop Voicing (p-word juncture rules on the Clitic Group domain; see Nespor and Vogel ibid: 158-159). These are either regressive or mutual consonantal assimilations in $C_1v_1C_2V_2$ where C_2 appears to be required to occupy a weak phonological position for the process to take place. This is only possible across words: while word-internally C_2 is strong since V_2 necessarily governs v_1 to silence it, in the cross-word situation v_1, a FEN, is taken care of by the FEN parameter, and consequently V_2's government hits C_2 weakening it, due to the absence (or 'extraprosodicity' – see the

[10] Cf. Scheer (2001).
[11] How Raddoppiamento Sintattico is stopped from applying across Phonological Phrase boundaries is not clear.

next section) of the boundary-marker of the second word. It is unclear why such conditions might be imposed upon the application of assimilation in certain languages, and a better understanding of assimilation as such in CV phonology is also called for.

Even more embarrassing are, however, cases of word-final voicing as in Sanskrit, where it is a U-level p-word juncture assimilatory process which is problematic for most phonological theories anyway since it is triggered by vowels and sonorant consonants, too. The examples in (12a) come from Nespor and Vogel (ibid: 118, 230) and illustrate Sanskrit, and exactly the same happens in Slovak (12b) (cf., e.g., Blaho 2003, 2004; data from Blaho 2004: 46).

(12) a. *sat – aha* *sad – aha* 'good day'
samyak uktam *samyag uktam* 'spoken correctly'
tat namas *tad namas* 'that homage'

b. *vojak* [k] 'soldier Nom.Sg.' *vojaka* [k] 'soldier Gen.Sg.'
vojak ide [g] 'the soldier goes'
les [s] 'forest Nom.Sg.' *lese* [s] 'forest Loc.Sg.' *les je* [z] 'the forest is'

Notice that these cases are puzzling since the presence of a local source for voicing assimilation is questionable in vowels and sonorant consonants, and the analysis sketched out in the previous paragraph to describe cross-word assimilation is even less probable here, the trigger not always being a consonant at all. What is still worth mentioning, though, is that mysteriously enough, both of these languages happen to be also characterized by word-final devoicing (and, in Sanskrit, deaspiration too). This means that FENs in these languages possess such a diminished licensing power that their C's are unable to support any laryngeal features, although they seem to desperately seek for one, even if it is shared by a following segment, be it a consonant or a vowel. This still does not account for where that laryngeal feature comes from in the case of sonorant consonants and vowels, but it is evident that the process takes place at the level of the melodies since supramelodic affiliation plays no role. It follows, then, that the presence or absence of the boundary-marker plays no role either, therefore the theory

predicts that rules like these cannot be bounded by it and as a consequence they must be U-level rules. Although this prediction still needs empirical justification, it is borne out at least in the two examples mentioned above, i.e., in Sanskrit and Slovak. In addition, it is also predicted that cross-word voicing always goes hand in hand with word-final devoicing and/or deaspiration, and the strict/permissive typology is not expected to influence the appearance of such a rule in a language – further predictions to be checked against a larger set of data.

4.3 How does the boundary-marker work?

Connected speech has not been given much attention in CV phonology. Tobias Scheer (2001, 2004 and to appear) has looked into how morphological information is projected into phonology and found that, in contrast to previous, diacritical and Lexical Phonological accounts, CV phonology armed with the boundary-marker offers a representational, privative and non-procedural alternative. Representing morphosyntactic boundaries with phonological material improves the model since it is mysterious why phonology should be sensitive to non-phonological information unless that information is "translated" into the "language" phonology speaks. Also, in a maximally constrained theory that information is privative: contrasts are expressed through the presence versus absence of some boundary signal rather than through its different values (e.g., plus or minus). Another advantage of the CV analysis lies in its predictive power: as we have also seen above, there is a causal relation between the phonological identity of the boundary-marker (i.e., empty cv unit) and its effects (i.e., triggering lengthening/gemination or blocking rules that require adjacency). In addition, Scheer suggests that the boundary-marker is not present in the lexicon but is inserted by the morpho-syntax. The insertion is governed by a simple parameter: in certain languages it applies on the edge of the utterance only, and at all word boundaries in others. This predicts, on the one hand, that phonological rules can apply within the

utterance or the word only. On the other hand, it predicts that within a language, either all the rules are bounded by the utterance or all of them are bounded by the word.[12]

Recall the findings of PP and notice that the picture is not that simple: constituents between the word and the utterance may also be designated as domains. In addition, in the same utterance boundaries of the same type may block the application of one rule but let go another. Thus we are forced to hypothesize, against Scheer, that, at least in strict languages, all the boundary-markers are present in the representation, but their fate is determined by the phonology: certain phonological rules will arbitrarily decide to ignore them and treat them as a kind of extraprosodic[13] skeletal material. The boundary-markers not ignored by a given rule will delimit its domain by blocking its application. Besides accounting for a wider set of empirical observations (i.e., the fact that all prosodic constituents between the word and the utterance may serve as rule domains), this analysis is capable of expressing the relevant parametric variation among languages (see Section 4.5).

The basic difference, then, is that while for Scheer the presence or absence of the boundary-marker depends on the arbitrary choice of the given *language* only, the present analysis claims it is rather an arbitrary choice of the phonological *rule*. Therefore, phonology has more control over morphological information: it is phonology that decides what to do with it rather than vice versa. Later, in Sections 4.4.1 and 4.5, we will see that Scheer's parameter does in fact exist, although it is not as general as he believes – instead, it seems to apply to the interplay of domains, phonotactics, and boundary-marking. In Section 4.5, it is shown how the constellations of parameter settings produce the distinction between strict and permissive languages and between the phonological strength of vowel- and consonant-initial words.

[12] Also, Scheer claims that in a permissive language like Moroccan Arabic the boundary-marker is never inserted, which means phonological rules are never blocked. This still needs to be verified, but see Section 4.6.

[13] This use of the term *extraprosodicity* is slightly different from its traditional sense: here, I am not talking about extraprosodic melody, but extraprosodic prosody. Extraprosodicity (or extrasyllabicity or extrametricality) in the traditional sense is of course rejected in GP, and so is here.

The chart in (13) compares the insertion[14] (13a) and the extraprosodicity (13b) analyses of two connected speech phenomena in an American English dialect (descibed in, e.g., Nespor and Vogel 1986). In this dialect, syllable-initial /l/'s are clear (e.g., in *lap, Yolanda, vigilant* – cf. Nespor and Vogel 1986: 65), whereas all word-final /l/'s are dark irrespective of the following segment (*pal, vigil, call, call Andy*)[15], which means that the domain of application of /l/-allophony is the word, and the boundary-marker of the following word blocks it. In the theory developed here, the rule of /l/-allophony makes /l/'s which are licensed and/or governed[16] by a following nonempty vowel clear (i.e., alveolar unvelarized), whereas /l/'s remaining unlicensed due to the inertness of the following empty v (as in word-final and preconsonantal positions) get velarised (dark). The rule of /t/-allophony, on the other hand, appears to operate in a different fashion. Recall from Chapter 3 that word-final /t/'s, although glottalised in isolation and preconsonantally (when they are followed by an empty v), change to a flap when followed by a vowel-initial word (up to the end of the utterance), which suggests that the government emanating from the vowel can hit the /t/ and cause it to lenite. (For a more detailed description of the analysis, see Section 4.4.2 and the following chapter.)

[14] Rejecting Scheer's analysis, which relies exclusively on morphosyntactic insertion, does not mean denying the possibility of morphosyntactic insertion as such. Clearly, morpheme and word boundaries are only known to morphology and syntax, and as the present theory claims, they can only be considered by phonology if they are "translated" into phonological material, i.e., into empty skeletal positions. Therefore, Scheer is right to say that it is morphosyntax that inserts the boundary-marker; however, as is evident from the discussion above, it inserts *all* boundary-markers, which are then evaluated by the phonological rules.

[15] In the same dialect, l̲'s remain dark when followed by a Class II suffix as in *vigilish*. The distinction between Class I and Class II affixation is beyond the scope of the present discussion.

[16] This uncertainty about the choice of the right term simply reflects the fact that no phonetic distinction is made in English between a licensed /l/ and a governed /l/ (and, as a third theoretical option, a licensed *and* governed /l/) – cf. the discussion in Section 4.4.2.

(13) American English /l/-darkening and /t/-flapping[17]

	call = *call Anita* blocking effect of boundary-marker	*hit* vs. *hit Anita* no (or different – see Section 4.4.2) blocking effect
(a) insertion	⇨ boundary-marker inserted	⇨ not inserted
(b) "extra-prosodicity"	⇨ it is present and has a blocking effect	⇨ it is present but flapping ignores it

In Scheer's framework (13a), the boundary-marker is expected to be projected to phonology in a phrase like *call Anita*, since its blocking effect is observed, as opposed to a phrase like *hit Anita*, where the status of the word-final consonant appears to have changed due to a following vowel, so the boundary-marker is not assumed to be active. Therefore, Scheer predicts that in identical syntactic positions, the boundary-marker may be present or absent depending on the quality of the final segment of the first word.

In contrast, in the "extraprosodicity" approach (13b), the boundary-marker is analysed as present and active in *call Anita* (since the domain of /l/-darkening is the p-word so no occurrences of the empty cv can be ignored by it), but it is analysed as present but inactive (invisible, extraprosodic) in *hit Anita*, the domain of application of t-flapping being the utterance, thus all boundary-markers but utterance-initial ones will go unnoticed.

Notice that this choice of the extraprosodicity analysis over the one involving boundary insertion draws heavily on Nespor and Vogel's (1986: 65) insight that the difference between the status of the word-final /t/ in *caught Andy* (intervocalic and flapped) and the /l/ in *call Andy* (word-final and dark) cannot be one of (re-) syllabification, but it is connected to the domains within which the rules apply.

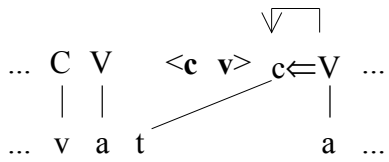
* * *

To illustrate the extraprosodicity and the blocking effect of the boundary-marker, let me present a possible analysis of French liaison, in which the floating melody of certain word-final consonants attaches to the onset of the following vowel-initial word.

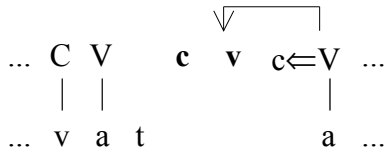
[17] Recall that /t/-tapping/flapping is part of the wider phenomenon of /t/-allophony also including glottalisation and aspiration. See Chapter 3 for the data and the following chapter for the analysis.

Apparently, this consonant can only land in a C position where it is governed (and licensed), otherwise it remains floating and therefore unpronounced[18]. The partial representation of *un* [*savant*]$_A$ [*anglais*]$_N$ in (14a) shows how the floating melody of the final consonant of the adjective, symbolised by /t/, docks onto the first c position of the following word when the boundary-marker is extraprosodic, i.e., invisible for the rule of liaison. However, when the boundary-marker serves as the boundary of the domain of application for liaison, as in (14b), it prevents the floating melody from finding a suitable landing site by absorbing the government emanating from the word-initial vowel. Thus the /t/ will remain phonetically uninterpreted, that is, 'final consonant elision' takes place.

(14) a. liaison: *un* [*savant*]$_A$ [*anglais*]$_N$ 'a learned Englishman'



b. no liaison: *un* [*savant*]$_N$ [*anglais*]$_A$ 'an English scientist'



In (15) the situation when the following word begins with a consonant is exemplified. Notice that in that case the presence (15a) or absence (15b) of the boundary-marker makes no difference: no governed empty c position is accessible for the floating melody in either way.

[18] As noted in Szigetvári (1999: 62, fn. 68), it is being governed, rather than being licensed, which triggers the interpretation of the lexically floating melody, although his analysis is slighly different since he works with a VCVC skeleton, and ignores the variation due to prosodic structure.

(15) no liaison: *un* [*savant*]$_A$ [*garçon*]$_N$ 'a learned boy'

a.

```
                    √
... C  V     c   v  C⇐V  ...
    |  |            |  |
... v  a  t         g  a  ...
```

b.

```
                    √
... C  V    <c  v>  C⇐V  ...
    |  |            |  |
... v  a  t         g  a  ...
```

The difference lies in the status of the word-initial consonant of the second word. While in (15a) it is licensed only, since the government charge of the following vowel is distracted from it by the boundary-marker, in (15b) it is governed *and* licensed. Although, as we have seen above, the boundary-marker is extraprosodic for liaison, and this supports the representation in (15b), that representation also predicts that word-initial consonants are weak and subject to lenition within the Phonological Phrase. I know of no process in French that affects such word-initial consonants, which may be indicative that, as in English, consonants at the beginning of the word are always protected by the boundary-marker as illustrated in (15a). For a more detailed discussion of "Scheer's parameter", see Section 4.4.1 and 4.5.

A phenomenon parallel to liaison in (14) is illustrated by Scheer's (2001) analysis of French gliding. As the data in (16) show, hiatuses are broken up by glide insertion (resulting from the vocalic melody spreading from the left) across a suffix boundary but not between a prefix and its base.

(16)	*lier* 'tie'	vs.	*bi-annuel* [biannyɛl] 'biannual'
	liais [lijɛ] 'I tie'		*anti-existentiel* [antiɛgzistasjɛl] 'anti-existential'
	lions [lijɔ] 'we tie'		
	lia [lija] 'I tied'		

Scheer claims that the difference is due to the fact that morphology does not project a boundary-marker between the root and the suffix (in, e.g., *lia* – 17a) but it does so between the prefix and the root (*biannuel* – 17b). As a consequence, the vowels in hiatus are adjacent and the intervening c position is intervocalic in (a) but not in (b).

```
(17) a.  C  V  c  V          b.  C  V  c  v  c  V  ...
         |  |     |              |  |              |
         l  i     a              b  i              a  nnuel
```

In PP, French gliding would be classified as a p-word-level rule, with the proviso that prefixes form separate p-words or at least are adjoined to p-words. Accordingly, the boundary-marker is not extraprosodic in (17b). Notice the similarity between (17) and (14) above: in both cases, there is an empty c potentially governed by a following vowel in (a), as opposed to (b), where government must silence the boundary-marker and thus no governed c is available for spreading.

It is worth mentioning at this point that the boundary-marker, and in fact empty c and v positions in general, will not block superficial, phonetic resyllabification, reported to take place in Romance languages[19] (e.g., in Nespor and Vogel 1986: 67ff). In Italian or Spanish, for instance, word-final consonants are claimed to resyllabify into the following vowel-initial word, which results in certain sequences containing and not containing a word-boundary being indistinguishable, as illustrated in (18) (data from Nespor and Vogel 1986: 68).

[19] In French, this phenomenon is frequently referred to as Enchaînement.

(18) a. Italian: *adocchio* '(I) sight' = *ad occhio (nudo)* 'with (the bare) eye'
b. Spanish: *alelado* 'stupified' = *al helado* 'to the icecream'

Admitting into phonological theory a resyllabification rule producing cross-word syllabifications identical to word-internal ones, however, leads to a pressure to duplicate syllabification rules, as is done in Nespor and Vogel (ibid.), syllabification proper taking place at the word level and resyllabification at a higher level, with most (if not all) phonological rules applying inbetween. This is needed since most (if not all) syllable-sensitive alternations fail to reapply when resyllabification modifies prosodic structure.

As an illustrative example, consider the case of s-aspiration in Spanish, whereby /s/ lenites to /h/ in "syllable codas" (i.e., when dominated by an unlicensed C position). Although in connected speech word-final consonants resyllabify into the following vowel-initial word phonetically, they do not usually do so phonologically (but see more on this issue in Section 4.4.2), i.e., lenited consonants do not strengthen back across words: *tienes espacio* 'do you have room?' is pronounced *tiene*.[h] *e*[h].*pacio* [20]. In French, Closed Syllable Adjustment replaces both /e/ and /ə/ by /ɛ/ in closed stressed syllables, even after phonetic resyllabification. Thus *première* 'first (fem.)' is /prəmjɛr/ in isolation as well as in *la première année* 'the first year' (cf. the masculine form *premier* /prəmje/), although native speakers syllabify the final /r/ into the onset of the following vowel-initial word (Nespor and Vogel 1986: 71). As the examples clearly show, phonology is sensitive to word-level syllabification only, and resyllabification (itself a theoretical outcast in most contemporary frameworks) cannot be accounted for without recourse to rule ordering, unavailable in Government theories. Thus, the cross-word resyllabification observed in Romance languages will be considered as a low-level phonetic phenomenon and henceforth be ignored. To put it differently, even though word boundaries may fade away phonetically, there is some evidence for the existence of empty skeletal material on which phonology applies.

[20] Cf. Kenstowitz (1994: 281).

Another remark in order here concerns the relationship between phonotactics and postlexical phonology. In CV phonology, the same mechanisms (viz., government and licensing) are claimed to be responsible for the definition of morpheme well-formedness (in the form of the constraints on possible consonant clusters, for example) as well as for post-lexical (assimilation, lenition and fortition) phenomena. It is expected then, that phonotactics is treated in the same way as any other rule: it may choose to consider or ignore the boundary-marker, but this time this choice defines two language types. If boundary-markers are visible to phonotactics, the language belongs to the type we call strict. Even phonotactics can be shown to apply within domains, however. In most strict languages all the boundary-markers are checked by phonotactics, so no initial RT-clusters are attested and FENs do not change their status (e.g. English) – the domain of application for phonotactics is the word. In some, on the other hand, only utterance-initial cv-units block the creation of falling-sonority clusters and only utterance-final FENs are silenced by parametric FEN-government – the domain of application for phonotactics is the utterance (see the discussion of Scheer's (2001) Bielorussian data in Section 4.5 for more detail). However, if phonotactics never takes the boundary-markers into account, then the language is permissive (in such languages, e.g. Greek, initial RT-clusters are well-formed at the beginning of any prosodic constituent including the utterance and the word). This view already suggests that the present discussion assumes the presence of the boundary-marker in permissive languages, too (see Section 4.6).

4.4 A typology of the effects of the extraprosodic boundary-marker

In this section we take all the combinatorial possibilities of consonant/vowel-final and consonant/vowel-initial words in connected speech, and make a comparison of cross-word and word-internal configurations in strict languages in cases when the boundary-marker is invisible to phonological rules. Throughout the discussion, we witness that

CV phonology makes predictions as to the cross-linguistic manifestation of edge effects which are all borne out by universal tendencies.

4.4.1 Consonant-initial words

In the table in (19), the four possible combinations of words followed by a (phonetically) consonant-initial word are sketched out. The first row shows the surface configurations, with traditional # standing for the morphosyntactic word boundary, and assuming that the segment occupying the underlined position is affected by some phonological rule ignoring the boundary-marker. The second row illustrates the patterning of empty and nonempty positions on the CV-tier in each case. At the bottom, for ease of comparison, the corresponding word-internal structures are given.

(19) Consonant-initial words:

a. V#C	b. V#C	c. C#C	d. C#C
... CV <cv>CV ...	... CV <cv>CV ...	... Cv <cv>CV ...	... Cv <cv>CV ...
...CVCV...		...CvCV...	

The fact that CV phonology predicts the cross-word and word-medial situations to be identical in all these cases is borne out by the data. For example, (19b), an intervocalic consonant affected by the phonology in the same way in both cross-word word-initial and word-internal positions is illustrated by Italian intervocalic spirantisation (Nespor and Vogel 1986: 209), whereby all the underlined /tʃ/'s (spelt <c>) in the following example sentence turn into /ʃ/, irrespective of whether at the beginning or in the middle of a word: *Il mio criceto cerca il suo cibo negli angoli della gabbia* 'My hamster looks for its food in the corners of the cage'.[21]

[21] The voiced affricate /dʒ/ undergoes spirantisation in the same way (and becomes /ʒ/), and the domain of application of Italian intervocalic spirantisation is the Intonational Phrase.

Hungarian regressive voicing assimilation exemplifies the configuration in (19c): it exists as a static phonotactic constraint (*zk morpheme-internally), and it applies across morpheme (*tízkor* 'at ten' with /sk/) as well as word boundaries (*tíz kör* 'ten circles' with /sk/).

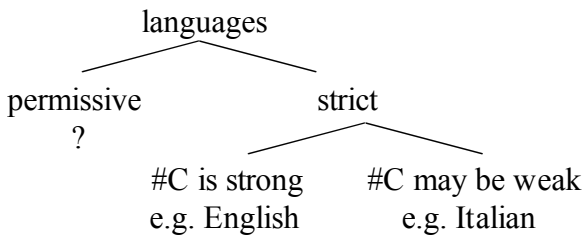
In sum, in all the situations in (19), it is correctly predicted that the cross-linguistic tendency is for the extraprosodicity of the boundary-marker to create the same picture as there is word-internally.

In addition, however, a parameter reveals itself. In certain languages, e.g. English, the word-initial consonant (of lexical words) will always be in a strong phonological position (i.e., licensed but ungoverned), as opposed to other languages, e.g. Italian (cf. the rule of intervocalic spirantisation, described above), with word-initial consonants changing shape post-lexically, which suggests that in languages of the English type the boundary-marker resists extraprosodicity in the case of consonant-initial words – an observation whose true nature is still unclear, but obviously this distinction is independent of the strict/permissive dichotomy mentioned above, both English and Italian belonging to the strict type.[22] Notice that this is exactly the parameter which Scheer (2001 and p.c.) describes, and which is mentioned at the beginning of Section 4.3 above. While Scheer claims the parameter generally governs the behaviour of all words, it is proposed here that it has a much narrower scope and only refers to consonant-initial words. As is argued in the next section, vowel-initial words enjoy a special status, and in languages like English certain rules *may* ignore their boundary-markers. Recall that we have argued throughout Chapter 3 and also in (13) above that word-final t's *are* influenced by a following vowel-initial word in GA, and it will be shown below that /l/-allophony (see (23) below) and /r/-allophony (25) in RP work in the same way (also see the next chapter for "ambisyllabicity" in English). In addition, the so-called non-rhotic dialects of English exhibit the phenomenon of r-liaison, whose analysis is expected to be analogous to that of liaison in French (in (14) above). As a humble tribute, and in order to highlight the parallelism between this study and his

[22] This typological difference between Romance and Germanic languages has long been well-known, mentioned in, e.g., Ternes (1986).

findings, I hereby name the parameter referring to the inalterability of consonant-initial words in strict languages "Scheer's parameter". The full typology developed up to this point, then, is schematised in (20), but find more discussion of parameters in Section 4.5.

(20)



To decide whether the bifurcation is relevant to permissive languages, too, a closer inspection of the data would be needed. Permissive languages present problems to the theory because they notoriously lack lenition processes, thus it is extremely difficult to decide whether a given phonological position is strong or weak. Even in Greek, the main argument for it belonging to the permissive type comes from diachronic changes (cf. Seigneur-Froli 2004, to appear).

4.4.2 Vowel-initial words

Consider the table in (21), the vowel-initial equivalent of (19). (21a-b) show the two subtypes of cross-word hiatus: it is clear that again, CV phonology predicts total identity between the cross-word and word-internal situations, which is supported by plenty of data, at least for (21a). In cases of hiatus resolution via vowel deletion, for example, it has been shown that there is a general tendency for deleting the first vowel in all morphosyntactic environments (Casali 1997).

(21) Vowel-initial words:

a. V#V	b. V#V	c. C#V	d. C#V
... CV <cv>cV ...	... CV <cv>cV ...	... Cv <cv>cV ...	... Cv <cv>cV ...
...CVcV...		...CV...	

In (21c-d), however, even if the boundary-marker is extraprosodic, there remain some empty skeletal material between the full positions, and as a result, the cross-word configuration is not identical to the simple word-medial CV string. Unfortunately, I have only come across few examples of vowels undergoing a process such as in (21d), one of them being vowel centralisation in Nawuri and related languages (Casali 1997: 502). Here high vowels become central in interconsonantal position, in both closed and open syllables (in Strict CV Phonology, CVCv and CVCV, respectively, which suggests that the trigger is the two nonempty consonants sandwiching the vowel). What is of interest here is what happens to vowels at word edges. As Casali reports (unfortunately, without any examples), word-final vowels in the CV#C environment may be affected by the change in the same way as word-medial vowels (as predicted in (19a)) as opposed to word-initial vowels, i.e., C#VC, which never get centralised. This difference between (19a) and (21d) is quite unexpected in any framework except CV (and Classical Government) Phonology. In fact, Casali uses the Nawuri example to argue for an asymmetry existing between word-initial and noninitial positions – an observation which naturally follows from strict CV representations.

If we turn our attention to (21c), we discover a number of cases illustrating it, a close inspection of which leads to a three-way classification. *First*, the underlined C in (21c) may resyllabify completely into a licensed position and behave as any other onset. Recall that this is the situation which is straight against CV Phonology's predictions, which turns out to be a strength rather than a weakness of the theory since, as argued in Kenstowitz (1994: 281), there are very few examples of this kind; in fact, phonological resyllabification counts rather as an exception.[23] One example described by Kenstowitz comes from Spanish: a "coda" /r/ is trilled in emphatic speech in both word-internal and

[23] The reader is also referred to the discussion of phonetic resyllabification in Section 4.3.

word-final position (in CV phonological terms, when followed by an empty v, which cannot license it), may be trilled when followed by a consonant-initial word, but cannot be trilled before a vowel-initial one (cf. (22)). What is particularly intriguing here is that all those many other phonological rules of Spanish affecting coda consonants (e.g., s-aspiration, already referred to in Section 4.3. above, or n-velarisation) apply differently, so this pattern seems to be the odd one out even within the system of Spanish.[24]

(22)	*martes* 'Tuesday', *mar* 'sea'	trilled
	mar verde 'green sea'	free variation
	mar azul 'blue sea'	no alternation possible: never trilled

Another example is /l/-darkening in certain dialects of English, e.g. RP, whereby coda l's become velarised, as in (23a-b), with the exception of word-final l's followed by a vowel-initial word (or suffix), which are pronounced as clear as their word-internal onset peers (23c).

(23) Clear and dark /l/'s in RP[25]

a. clear /l/'s: *leap, sleep, fellow, mylord*

b. dark /l/'s: *spell, spelt, shelter*

c. clear /l/'s: *spell it, call Ann, spelling*

It will be argued below that no convincing evidence has been found that these consonants do in fact resyllabify completely rather than take an intermediate position (traditionally referred to as ambisyllabicity).

The *second* strategy that a word-final C may follow is remain a phonological coda, e.g., in the case of Spanish s-aspiration referred to above, or l-darkening in certain American English dialects exemplifed in (13). In these cases we claim that the word-

[24] The relevance of the data, however, is doubtful since, as the audience of a doctoral research seminar pointed out to me, the phenomenon is much more complex. Notice that the more questionable the case the less likely it is to exemplify phonological resyllabification – a favourable development for the present theory.

[25] For more on the dialectal variation in /l/-allophony in English, see the next chapter.

boundary represented by the boundary-marker functions as a blockage for these rules, the (prosodic) word being the domain of rule application – an arbitrary feature of the rules themselves. Notice, then, that this is not a true subcase of (21c) since it is not analysed with the boundary-marker extraprosodic, and it has also been shown that it is a mere phonetic illusion that the final consonant has resyllabified into the following word (cf. Section 4.3).

Thirdly, the C may behave as neither an onset nor a coda but take a third form as if it was both (or neither): it is "ambisyllabic". English readily illustrates this pattern, containing at least two rules where the cross-word realisation of a consonant differs from both the coda and the word-medial onset. One is the distribution of GA /t/-allophones (see Chapter 3), whose basic characteristics can be briefly summarized for convenience as follows. Within words, an onset /t/ is flapped if it is followed by an unstressed vowel, but aspirated if it is followed by a stressed one (24a). Word-final /t/'s are glottalised pre-pausally and pre-consonantally (24b), but flapped if the next word starts with a vowel, irrespective of whether or not that vowel is stressed (24c). The point is that the cross-word allophone in C#V is different from the word-medial one (in being stress-insensitive), correctly predicted by CV phonology (for the analysis, see the next chapter).

(24) General American /t/-allophones

a. á[ɾ]om *atom*, a[tʰ]ómic *atomic*

b. hi[tˀ] me *hit me*

c. hi[ɾ] Ánn *hit Ann*, hi[ɾ] Aníta *hit Anita*

Exactly the same happens in (conservative) RP /r/-allophony. /r/ undergoes tapping/flapping, with an output identical to that of /t/-flapping, intervocalically, whenever it is followed by an unstressed vowel word-internally (compare (25a-b)), or any vowel across words (25c).

(25) RP /r/-allophony

a. [ɾ]: *courage, very, sorry, baron, laurel*

b. [r]: *courageous, reduce, red, bright, Henry, walrus*

c. [ɾ]: *for exámple, for ínstance, the other énd*

These examples illustrate the fact that the situation C#V is special and calls for a theoretical equivalent of cross-word ambisyllabicity. A possible analysis in Strict CV phonology will be elaborated on in Chapter 5. The basic idea is the difference in adjacency between the prosody and the melody in a theory operating with a host of empty skeletal positions. A word-medial consonant (26a) is adjacent to the following vowel both melodically (indicated by the simple arrow) and on the CV-tier (where the V licenses the C, indicated by the double arrow). Across words (26b-c), the two are only adjacent on the melodic tier, where the melody of the vowel (β in (26b)) can have an effect on the consonant. This effect, I claim, is government, which is not consumed by the boundary-marker when it is extraprosodic. The same word-final consonant, however, will remain uninfluenced by the following word when the boundary-marker is not extraprosodic, and the government emanating from the full vowel is needed to license and silence its vocalic position (26c).

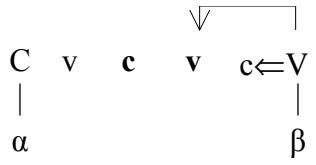
(26) a. ...CV...

```
c⇐V
|  |
α< β
```

b. ...Cv<**cv**>cV...

```
C   v   <c   v>   c⇐V
|                  |
α <————————        β
```

c. ...Cv**c**v**c**V...



Thus, three possible combinations of skeletal and melodic adjacency are possible, and the English cases (/t/-flapping (24), /r/-tapping (25)) are distinguished from the other examples cited above by the rules applying in three different ways accordingly. The question is whether the "resyllabifying" rules described above (RP /l/-darkening (23), Spanish trilled /r/ (22)) are essentially any different. It may simply be the case that, quite unexpectedly and exceptionally, there is no phonetic difference between the realisations of these consonants in situations (26a) and (26b), and that is why the superficial impression is that they have become onsets.[26] The exact definition of the structural descriptions of phonological rules may be of key importance: sometimes it is not the interaction of government *and* licensing which produces a given allophone, but reference to only one of the two antagonistic forces may prove to be enough for a change. A governed /t/ will be flapped in GA, a governed /r/ will be tapped and a governed /l/ will be clear in RP, a governed /r/ will be plain (rather than trilled) in Spanish, irrespective of the rest of the structure (i. e., of whether licensing applies, too). Recall the discussion of French liaison (14) above: it has been shown that the appearance of the word-final floating consonantal material is sensitive to government and ignorant of licensing relations.

This treatment of cross-word ambisyllabicity also predicts that, should the status of a word-final consonant get altered across a word boundary, it always changes into a (nother) weak allophone and never into a strong one, that is, there are no cases of cross-word "upgrading". This prediction follows from the claim that while vowel-to-consonant government takes place between melodies, and is thus capable of skipping empty skeletal positions, licensing, the force responsible for stability or strengthening, is

[26] Following a suggestion by Rubach (1996), the next chapter considers cross-word clear /l/ as ambisyllabic.

still considered to be contracted on the CV-tier[27]. Consequently, the licensing emanating from a word-initial vowel never reaches the preceding word because it "bumps into" the vowel's own (empty) c position, compare (23a) and (23b-c). Recall from the discussion above that the examples in which final codas seem to resyllabify as onsets phonologically (i.e., Spanish r-trilling and RP l-darkening) are either dubious or easily reanalysable as cases of ambisyllabicity. Therefore, CV phonology has again been shown to make the right predictions.[28]

Before we move on to discuss the relevant parameters, another remark is in order here. As early as in Section 1.1.1, the phonological Projection Principle was introduced and its repeated violations were sketched out both in Classical GP and CV phonology. We saw there that most of the violations stem from the insufficiency of how these analyses attempt to treat cross-word phonology. Notice that the account outlined above and detailed in the next chapter does not delete either prosodic structure or governing/licensing relations – in full conformity with the Projection Principle.

4.5 Language typology and parameters

4.5.1 Word-initial consonant clusters

In the discussion above several parameters with the ensuing language typology were touched upon. In addition to the FEN-parameter familiar from Standard GP (following Kaye 1990b), according to which languages have the choice of allowing or disallowing for domain-final empty nuclei (FENs), the cv-parameter was introduced in Section 4.1. In the original suggestion, i.e. Scheer (2001, 2004, to appear), it refers to the simple presence of the boundary-marker in strict languages and its absence in permissive systems. In Section 4.3 above, the cv-parameter was modified so as to refer to the

[27]But see Chapter 7.

[28] The next chapter examines the predictions of a VC skeleton (cf. Dienes and Szigetvári 1999, Szigetvári 1999b, and the discussion in Chapter 2), and concludes that they are erroneous in several respects; the point is illuminated with the examples of GA t-allophony.

operation of phonotactic constraints only. A summary of the phonotactic effects of the boundary-marker is given in the typology of word-initial clusters in (27) (adapted to our present purposes and theoretical vocabulary from Scheer to appear).

(27)	#CV	#TR	#RT	Parameter settings	Examples
a. No initial clusters	yes	no	no	phonotactics evaluates cv + IG switched off	Ticuna, Klamath, Finnish, Mokilese, etc.
b. Strict	yes	yes	no	phonotactics evaluates cv	typical IE (English, French, etc.)
c. Permissive	yes	yes	yes	phonotactics ignores cv	Modern Occidental Semitic, Berber, Greek, Slavic
d. #RT-only	yes	no	yes	??	unattested

Since single initial consonants are well-formed in all languages, the #CV-column contains yeses in all cells. Falling-sonority clusters are ruled out whenever the boundary-marker is visible to phonotactics (as in (27a-b)) because the ECP of the empty v in the boundary-marker is not satisfied (cf. (2a) above). In case Infrasegmental Government (IG) is available in order for certain consonant sequences to form a closed domain (as in (3a) above), a restricted set of initial two-consonant clusters appears, resulting in strict languages (27b). In Scheer (2001) it is claimed that permissive languages treat all initial clusters alike irrespective of their sonority profile or other characteristic, so IG is not an option in (27c). Most importantly, however, the system predicts the nonoccurrence of #RT-only languages (27d); notice that no parameter setting could produce such a situation.

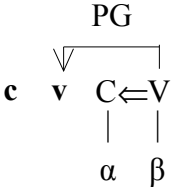
4.5.2 The strength of the left edge

The reader may also be able to recall from Section 4.1 that originally, another prediction of the presence vs. absence of the boundary-marker was emphasized, viz., the presence or absence of Coda Mirror (CM) effects word-initially. For convenience, the two CM environments, the word-initial (#_) and post-coda (C._) positions, are illustrated in (28a)

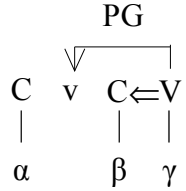
and (28b), respectively. CV phonology claims that both are strong since both C's are preceded by an empty nucleus which consumes the government of the following nonempty V.

(28)

a.



b.



If it is the boundary-marker that makes the word-initial position strong phonologically, the same environment must be weak (or at least behave in the same way as medial intervocalic consonants) when the boundary-marker is not there (or is extraprosodic), i.e., in permissive languages. This is because then the PG of (28a) hits the initial C itself.[29] As was hinted at earlier, in the languages classified by CV phonologists as permissive (i.e., Modern Occidental Afro-Asiatic and Slavic (Scheer 2001), and Greek (Seigneur-Froli 2004, to appear)), a difficulty arises in the distinction between weak and strong phonological positions as those languages do not exhibit lenition and fortition phenomena that would form the basis for the decision. Seigneur-Froli (2004, to appear) argues that the beginning of the word, as opposed to the post-coda position, is weak in Greek, unlike in Germanic and Romance, where both resist lenition. The data she cites, however, is diachronic, and exclusively refers to the evolution of these languages.[30] Although there is no denying the relevance of diachronic evidence in phonological theory, it is difficult to identify its real significance in synchronic argumentation. Moreover, even if some kind of static, out-of-context weakness is attested in permissive languages (or at least in Greek) which manifests itself in the diachronic career of words,

[29] A further prediction of CV phonology and CM theory is that while parametric variation is expected at the two edges (cf. the cv-parameter and the FEN-parameter), it is not found morpheme-internally, i.e., in (28b). The prediction is borne out by the data: no language is on record where word-initial consonants are strong, whereas post-coda consonants are weak.

[30] Even the title of Seigneur-Froli (2004) states explicitly that the evidence comes from *diachronic* consonant lenition.

it does not necessarily follow that the same behaviour is found synchronically.[31] It has been stressed several times above that even in strict languages the left edge sometimes becomes weak in connected speech, as in Italian and Spanish. In conclusion, the separation of the role of the boundary-marker in phonotactics and dynamic processes seems to be justified, and so does the denial of the equations "strict = strong word-initial C", "permissive = weak word-initial C".

4.5.3 Phonotactics

In the present thesis, then, as is proposed at the end of Section 4.3, the strict/permissive bifurcation is argued to stem from whether the boundary-marker is extraprosodic when phonotactic constraints are checked. By phonotactics, I mean another bundle of parameter settings including whether IG can be contracted, whether the ECP is suspended for the boundary-marker in exceptional *sC* (or other) clusters, whether FENs are able to govern and/or license, etc. The only difference between phonotactics and allophony-rules, then, lies in their content: phonotactic constraints are manifestations of language-specific parameter settings, whereas allophony-rules are language- and segment-, prime- or natural-class-specific spell-out statements about the phonetic interpretation of melodic primes in licensed, governed, unlicensed, and ungoverned positions. Although CM and CM+ (see Chapter 2) are able to make predictions as to general tendencies in these spell-out rules (i.e., strong (licensed ungoverned) positions support the melodic expression of the target, other positions count as weak, and there are two different weak positions (governed vs. unlicensed), in which two different types of lenition are expected (vocalic vs. consonantal, resp.)), it remains totally language-

[31] The same difference between diachronic and synchronic behaviour is exemplified by function words. Scheer (to appear) contends that no distinction need be drawn between major lexical and functional categories, since he bases his arguments on history: the left edge in IE is strong in both cases, cf. Latin *duo* 'two', *tres* 'three', *pro* 'for', *per* 'per', where no lenition of the initial plosives took place during the evolution to French *deux*, *trois*, *pour*, *par*. Thus they are found to be just as protected from weakening as the beginning of major categories. Nevertheless, in many languages such as English, unstressed function words cliticize onto the preceding word phonologically, that is they form a single p-word (or clitic group, depending on theoretical taste) with it. Should the relevant segmental conditions be met, the beginning of function words *does* lenite then, cf. *lie to me* and similar examples from Chapter 3.

specific how these universal tendencies are realized, i.e., whether there is a phonetic difference between strong and weak, governed and unlicensed, etc. allophones.[32]

Both phonotactics and "dynamic" rules, however, can be bounded within their domains. It has been shown above how arbitrarily phonological rules choose their domains of application. In addition, several authors, e.g. Booij (1999) claim that the p-word serves as the domain of phonotactic constraints: it is well-known that consonant clusters like /pb/, well-formed across p-words (e.g., in a hypothetical compound *cupboard* having the literal meaning 'a board for cups', or within a phrase like *keep beating*) are done away with within the limits of a p-word (in synthetic/monomorphemic *cupboard* 'a closet with shelves where dishes, utensils, or food, and not only cups, are kept'). The more general principle underlying the *cupboard*-case is the well-known fact that not all combinations of well-formed syllables yield a well-formed p-word, i.e., although both *cup* and *board* are existing (that is, licit) monosyllables, they cannot be combined into a single licit p-word without any modifications. Also, the attempt at joining the apparently well-formed right edge /kst/ of a syllable like *text* with the apparently well-formed left edge /str/ of a syllable like *strip* will result in the string /kststr/ unattested p-word-internally.[33] Moreover, the behaviour of word-internal clusters is generally determined by whether they occur word-initially, which made Kahn (1976), for instance, circumscribe the legal syllable-initial clusters of English with reference to the existing word-initial sequences. Medial *sC*- and *sCC*-clusters, e.g., are treated as onsets by stress rules in the great majority of the cases, even if they exhibit sonority falls, cf. the stress patterns of *mínister* and *índustry*, where the antepenultimate stress can only yield if we assume the syllabifications *mi.ni.ster* and *in.du.stry* with the underlined penults being light and therefore non-stressable[34]. By contrast, even if a cluster, say /kn/, contains rising sonority and is therefore well-formed in other Germanic languages like German (and also in Old and Middle English) in word- and syllable-

[32] As Harris (1994: 271) expresses the same idea in his Epilogue: "[...] it needs to be stressed that the emerging model stops short of being fully deterministic. It still possesses a certain degree of 'slack', permitting situations in which more than one derivational outcome is well-formed. This is as it should be."
[33] Observations like this have led a number of phonologists to describe the special status of word edges with reference to an obligatorily peripheral constituent (often referred to as the Appendix) which lies outside the scope of basic syllabification.
[34] Whatever this means in Strict CV/VC phonology.

initial positions, (Present-day) English fails to recognize it as an onset due to its illicitness word-initially, and thus the /k/ in *acne* is glottalized rather than aspirated.

We conclude, on the basis of like phenomena, that the domain of phonotactic constraints is the word, rather than the syllable, as is traditionally thought. This also explains why, even in languages which allow for branching onsets, no such clusters are created upon the concatenation of words: since the contraction of IG relationships is in control of phonotactics, and for phonotactics all boundary-markers are visible, therefore it follows that consonants refuse to form Scheer's "closed domain" even if other rules are capable of crossing word boundaries. This can be illustrated by English affricate-formation (cf. Section 3.1.4): when /r/ stands after /t/ or /d/ within the same word, it combines with them to produce affricate-like sequences: *train, drain, attract*, etc. Since this process is exclusive to 'branching onsets', that is, its prerequisite is the existence of Infrasegmental Government relation between the consonants in question, the fact that it does not apply across words (cf. the /tr/ sequences in *nitrate, petrol* vs. *night rate, hat rack*) is easily accounted for.

4.5.4 Bielorussian

Scheer (to appear) proposes that the boundary-marker, being a totally predictable object, is not part of the lexical representation of words, but, as has been mentioned above, is inserted by the morpho-syntax. Consequently, since phonology is sensitive to its presence or absence, all of phonology must apply *after* this morpho-syntactic insertion. That is, there is no distinction between lexical and post-lexical phonology: only post-lexical phonology exists, applying upon a flat, homogeneous CVCV-structure, which is enhanced with morphosyntactic information expressed through a unit that enjoys a truly phonological identity. In light of this, phonotactic constraints may only manifest themselves in the words' pronunciation in isolation, since they may become blurred in connected speech. Let us illuminate this point with Scheer's (2001, to appear) Bielorussian examples.

To lend support to his parameter, viz. that the boundary-marker is either inserted at the beginning of each word or at the beginning of the utterance only, Scheer examines strict languages in which phonology does not care about word boundaries: Central Italo-Romance and Bielorussian. Unfortunately, he does not illustrate connected speech in languages where the boundary-marker is inserted at the beginning of all words within the utterance, so his view of the English pattern is left unclear. Romance languages have been amply exemplified above (esp. Italian in Section 4.4.1; Sardinian has also received much attention lately – see, for instance, Ladd and Scobbie 2003), and Scheer pays more attention to Bielorussian, where he investigates the behaviour of /v/, which surfaces as [v] in intervocalic position (*koro**v**a* 'cow NOMsg') and the Coda Mirror (i.e., in #__V e.g. ***v**ada* 'water' and in postcoda e.g. *bar**v**a* 'coloration'). In codas, it is [w] (*koro**w**ka* 'little cow', *koro**w*** 'cow GENpl'), and utterance-initially before a consonant, it is [u] (***u**dava* 'widow'). Crucially, this latter statement only holds for the utterance-initial position, since utterance-internally a [w] appears after vowel-final words (*taja **w**dava* 'this widow' vs. *brat **u**davy* 'the widow's brother'). Initial prevocalic v̲'s, however, are stable (*taja **v**ada* 'this water', *brat **v**ady* 'the brother of the water'). Scheer finds that cross-word situations mirror the corresponding word-medial configuration as shown in (29), and concluded that in Bielorussian utterances are headed by a CV-unit (i.e., it is a strict language), within utterances, however, no CV-units occur. Bielorussian v̲-allophony, then, can be described as a U-level rule complex affecting the phonetic interpretation of a floating **U** element: attached to a licensed (and optionally also governed) C position, it finds itself in a strong phonological position and surfaces as [v]; in an ungoverned and unlicensed position it lenites to [w]; and sometimes it is forced to occupy a V position (see (30) below) when it is [u].

(29) a. brat **u**davy = **u**dava *empty v of FEN/cv is interpreted by **U** spreading*

b. brat **v**ady = bar**v**a */v/ is after an empty v and thus ungoverned*

c. taja **w**dava = koro**w**, koro**w**ka

/v/ is followed by an empty v and thus unlicensed

d. taja **v**ada = koro**v**a */v/ is intervocalic*

Now the reader is invited to verify the italicized analyses of (29) in the representations in (30) below.

(30)

a. ... C **v** C v C V ...
X
... t **U** d a ... *brat udavy*
U d a ... *udava*

b. ... C v C V ...
... r **U** a *barva*
... t **U** a ... *brat vady*

c. ... C V C v C V ...
... j a **U** d a ... *taja wdava*
... r o **U** *korow*
... r o **U** k a *korowka*

d. ... C V C⇦V C V
... j a **U** a d a *taja vada*
... r o **U** a *korova*

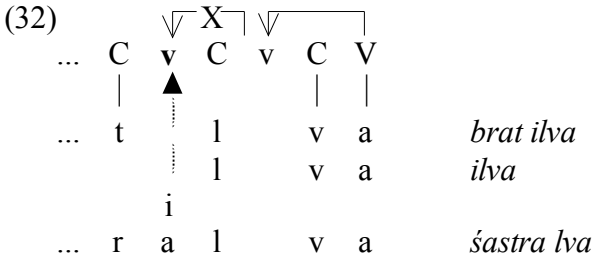
The other process that Scheer analyses is i-epenthesis. The relevant data is in (31).

(31) Bielorussian i-epenthesis

lew 'lion NOMsg'	brat **i**lva 'the brother of the lion'
ilva 'lion GENsg'	malady lew 'young lion'
tam jość lew 'there is the lion'	śastra lva 'the sister of the lion'

The exciting thing about this set of data is the fact that the word for 'lion' has a stem with a yer and no nonempty vowel. The yer is phonetically realized when followed by an empty nucleus unable to govern it (as in the nominative singular), but remains empty whenever there is a proper governor to the right (such as the *–a* in the genitive singular). Since Bielorussian is a strict language, the boundary-marker is there, which is not properly governed when the yer in the stem is, and therefore its empty v must be interpreted phonetically. Bielorussian opts for inserting an /i/ in such cases, and that is how the form *ilva* is derived. As the quality of the final segment of the preceding word in connected speech seems to have an effect on i-epenthesis (cf. *brat ilva* vs. *śastra lva*), Scheer concludes that FENs cease to get parametric silencing utterance-medially, and must receive the same epenthetic /i/ whenever they are trapped between two consonants and no governor is available. See the representations in (32).

(32)



Notice that Bielorussian only rejects #RT-clusters utterance-initially – utterance-internally such clusters do arise, as in *śastra lva*. Thus, the fact that Bielorussian is a strict language is only apparent from the citation form *ilva*.[35] The boundary-marker is

[35] The question of where the /i/ comes from remains unanswered in the analysis. Scheer (2004) rejects the language-specific interpretability of empty nuclei and argues that all alternating vowels other than /ɨ/ are underlyingly present but floating melodies. Accordingly, all surfacing i's which alternate with zero should

very easy to ignore, similarly to Spanish or Italian, where word-initial consonants lenite within the utterance. This is not a universal property of strict languages, though.

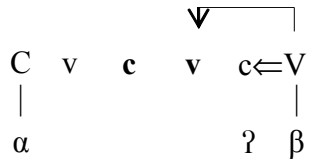
4.5.5 Consonant- vs. vowel-initial words

In Section 4.4.1, it was argued that a parameter ("Scheer's parameter") is responsible for English, as opposed to Italian, Spanish (or Bielorussian), initial consonants being always strong, i.e., for the boundary-marker always being visible. It is also claimed here, and will get even more emphasis in the next chapter, that vowel-initial words exhibit special behaviour with respect to the underparsing (or extraprosodicity) of the boundary-marker: at least in English (and most probably in French, too[36]), they are able to let phonological rules ignore it, resulting in the "ambisyllabicity" of preceding final consonants. Other languages, e.g. German, may be even stricter strict languages, and disallow for the extraprosodicity of the boundary-marker in vowel-initial words, too. In German, its constant presence results in the constant impression of hiatus word-initially, which in turn leads to hiatus-breaking by the insertion of the glottal stop. As a consequence, all words start with a consonant on the surface. Another consequence is the absence of liaison in German: e.g., although it possesses an r-dropping rule similar to the one in non-rhotic English (i.e., r's are only pronounceable – "licensed" in the traditional sense – in governed position), word-final r's are never reached by the government issued by an initial vowel since it is always taken up by the intervening boundary-marker. Thus, unlike in English, no linking occurs. Cf. (33).

originate from the linking of a floating **I** element. The drawback of this account is that it is forced to assume underlyingly floating **I** elements for boundary-markers (to derive *ilva*) and FENs (to derive *brat ilva*), too, or to the left of all stems containing a yer at the beginning similarly to *lew*.

[36] Evidenced by liaison phenomena.

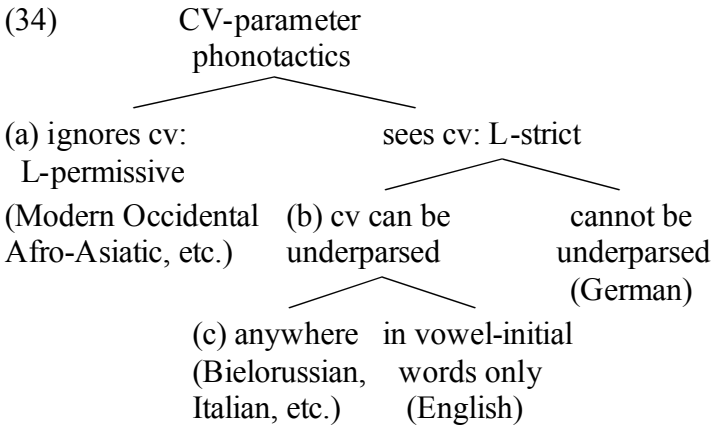
(33)



4.5.6 Summary of the cv-parameters

A summary of the parametric variation connected to the boundary-marker is given under (34). "L" stands for the left edge, and is used here to suggest that a parallel terminology can be applied to the right edge of words (see (37) below). (34c) is "Scheer's parameter".

(34)



4.5.7 FENs will be FENs?

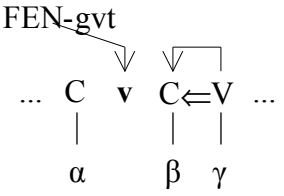
At the right edge of words, the empty position subject to parametric variation is the FEN, introduced in Chapter 1. Since Kaye (1990b), GP has posited empty nuclei whenever a word ends in a consonant on the surface. Languages in which this is an option are claimed to "license" FENs via parametric FEN-government. Henceforth such languages will be referred to as right-permissive, since they permit an empty category at

the right edge. Typical IE languages like English and German are R-permissive. However, if the FEN-parameter is switched off, FENs are not governed, and the language is R-strict in the sense that all words strictly end in a vowel, as in Italian.[37]

This section is, quite understandably, primarily concerned with the behaviour of FENs in connected speech, and the interaction of morphosyntax with the FEN-parameter. In Scheer's (to appear) conception, two types of morphosyntactic action affecting FENs are possible. First, morphosyntax is entitled to grant licensing or governing abilities to FENs. For example, if a FEN can govern, as in English, word-final falling-sonority clusters appear in the language, since the FEN will properly govern the empty nucleus sandwiched between the consonants. If a FEN can license, as in French, TR-clusters are well-formed word-finally, as the second consonant is endowed with the strength required to contract IG with the first consonant[38].

Second, in Scheer's view, upon the attachment of a suffix morphosyntax may decide to allow the FEN of the preceding morpheme to be unpronounced. As a result, the initial consonant of the attached suffix (if there is one) is "demoted" from strong CM to weak intervocalic position. As is illustrated in (35), the government of the nonempty vowel hits the preceding consonant instead of the empty v, which is taken care of by morphosyntax.

(35) FEN-gvt



It is evident that this analysis can only be adapted to the cross-word situation in cases when the boundary-marker is not inserted, as Scheer would put it, or, using our terminology, when it is extraprosodic. Moreover, it assumes that FENs are only normally governed utterance-finally. Within the utterance between morphemes, FENs

[37] For a full typology of the parametric setting of FENs, see Scheer (2004: 651).
[38] Governing relations are head-final for Scheer.

will not be FENs any more, and either special morphosyntactic order is needed for them to survive as in (35), or else phonology must take the stage. In the Bielorussian examples above, former FENs receive PG (see (30b)) or surface phonetically ((30a) and (32)).

This is in sharp contrast with the Standard GP treatment of FENs. In Kaye (1995), for instance, all rightmost analytic domain boundaries allow for the inaudibility of an empty nucleus. In English compounds, e.g., [[black][board]], there is a parametrically silenced FEN at the end of both *black* and *board*. The point is the fact that the same mechanism is responsible for the emptiness of the FEN in *black* when pronounced in isolation, and when it is the first term of the compound. After all, this is the motivation for talking about *domain-final* rather than *word-final* empty nuclei. Therefore, FENs will be FENs even after their host has been concatenated with another morpheme.

As an illustration, consider the two derivations of the famous English word *sixths*. Kaye claims that it has an intricate domain structure, [[[sɪks**v**]θ**v**]s**v**], and each of the empty v's are "licensed" by virtue of being domain-final. As opposed to this, Scheer claims that the morphological boundary preceding them entitles them to govern the empty v to the left. That is, the empty v's govern each other in a chain, and only the word-final v is a parametrically legalized FEN.

Across words, then, only utterance-final FENs exist. Notice that although Scheer's analysis is right for Bielorussian, it is wrong to claim that this treatment of FENs is universal. It is wrong since if that was the case, former FENs would be forced to be pronounced across words in languages where the boundary-markers are present, as in (36a). It is wrong since if that was the case, the government emanating from a word-initial vowel would never reach a preceding word-final consonant, either because it is consumed by the boundary-marker (36a) or by the FEN of the preceding word (36b). Recall from the discussion in Section 4.4.2 above that, at least in certain languages like English, such vowels *are* able to exert some influence.

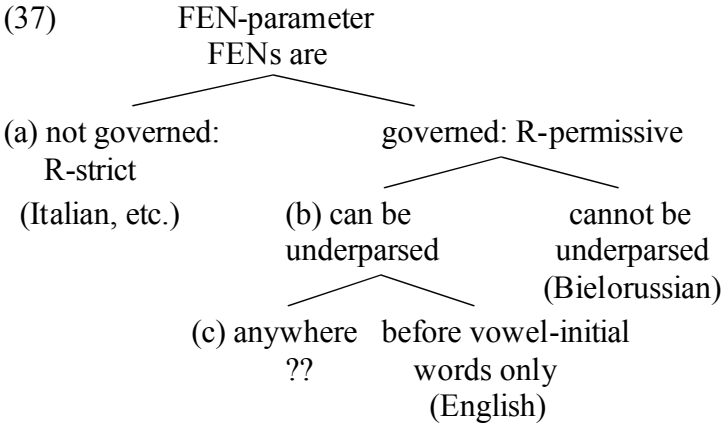
(36)

```
a.                  b.
C  v  c  v  C⇐V     C  v  <c  v>  C⇐V
|           | |     |             | |
α           β γ     α             β γ
```

In addition, at the end of Section 4.2, the analysis of cross-word assimilations inoperative morpheme-internally made crucial use of the reference to FEN-government present in the representation. If that analysis is on the right track, this constitutes another argument for FENs remaining FENs in certain cross-word configurations.

In sum, while I accept Scheer's treatment of Bielorussian and similar languages, I insist that sometimes FENs keep their original status, i.e., morphosyntax detects a domain boundary and informs phonology about it – as long as the FEN-parameter is on in the language, FEN-government silences the empty v next to that boundary. When exactly this happens is not entirely clear, though. No reliable conclusion can be reached until an extensive cross-linguistic study is carried out. I suspect that the variation is parametric again, similarly to the distribution of the boundary-marker in (34) above, and a parallel classification may be possible, as illustrated in (37).

(37)



Accordingly, even R-permissive languages have the choice of where they allow for the parametric government of empty nuclei, and where they force phonology to take care of

them in the same way as their internal peers. If FENs are only licit utterance-finally, as in Bielorussian, then this means that no internal empty v's are invisible to phonology, and the utterance behaves as one big (synthetic) domain (37b). A look at English reveals that a(nother) strange difference may exist between consonant- and vowel-initial words (37c): as has been shown above and will be utilized in the next chapter, vowel-initial words are able to ignore FENs as if they were not there (or, as if FEN-government was in effect) and govern the final consonant of the preceding word instead. Consonant-initial words, on the other hand, appear to consider preceding FENs as any other empty v's: C-to-C assimilations, e.g., are just as possible across word boundaries as within, which indicates that in a C_1vC_2V sequence the nonempty V governs the empty one in the same way irrespective of its origins (cf. (36b)). This may be connected to the requirement in English that all initial consonants be in a strong phonological position, violated in a configuration like (35) above. Finally, the system predicts the existence of a language type in which all FENs remain FENs (37c), but the verification of this claim is still subject to further study, although languages like Sanskrit and Greek, described above as having assimilatory processes exceptionally triggered by the word junction only, seem to be good candidates.

Therefore, boundary-markers (34) and FENs (37) might turn out to be driven by parallel parametric variations, which is not surprising regarding the fact that in one way or another both mark morphosyntactic boundaries. To decide whether possible implicational relationships can be identified between the two parameter families, a more extensive cross-linguistic study is needed.

4.6 Further issues

This chapter has argued that the boundary-marker is assumed to be present to the left of each (lexical) category, at least in so-called strict languages, and the phonological rules spelling out the realisation of sound segments contain information about which prosodic constituents serve as their domains of application. Boundary-markers delimiting those

constituents will block the rules, others will be skipped being extraprosodic in some sense. Although certain language-specific tendencies can be formulated with respect to what prosodic constituents serve as domains for the rules (e.g., Bielorussian seems to prefer the utterance, and even in English, most rules producing segmental alternations in most dialects also choose the utterance), we claim that phonological rules select their domains arbitrarily.

Formulating the syntax-phonology mapping algorithm, i.e., the formation of the constituents of the prosodic hierarchy, however, is beyond the scope of the present discussion, so we simply accept the PP view (e.g., that of Nespor and Vogel 1986). Nevertheless, it is worth stressing that making reference to the members of the prosodic hierarchy (in (9) above) does not entail that phonological representations are necessarily hierarchical above the CVCV-tier. Within the syntax-phonology interface, one of the most controversial issues has been how phonological rules can be made to refer to sentence structure, and two competing views have arisen. Although some, e.g. Nespor and Vogel (1986), Hayes (1989b), Zec and Inkelas (1990), Selkirk (1986), argue that phonological rules refer to various types of prosodic structures, which are themselves built by rules that refer directly to syntactic structures, others, e.g. Kaisse (1985), contend that phonological rules may directly refer to properties of surface syntactic representations, and even Nespor and Vogel (1986) and Selkirk (1986) admit that there are phonological rules which directly refer to syntactic conditioning. If the direct reference theories are right, then it is not impossible to imagine a simple morpho-syntactic algorithm not only inserting cv's to indicate morphosyntactic boundaries but also labelling them as to what constituents they delimit.

What is evident from the above discussion, however, is that constituents smaller than the word, i.e. the syllable and the foot, also featuring in the hierarchy in (9), do not play a role here. While rules traditionally analysed as bounded by the syllable make direct reference to surrounding empty and nonempty categories (as explained in Section 4.2), foot-level phenomena are thought to be governed by the stress-sensitivity parameter (see below and the following chapter). Additional support for expressing syllabic and foot-level constituency in terms different from the prosodic units is given

by the observation of their differences made by the proponents of PP themselves, and based on independent evidence: "... in light of many differences between metrical units and those which function as rule domains, a number of researchers have suggested that the two constituent types belong to separate hierarchies[39] (Selkirk 1986, Zec 1988, Inkelas 1989)", say Inkelas and Zec (1995: footnote 3).

Another topic exceeding the present discussion's scope is the examination of permissive languages. These are the languages which freely tolerate all types of consonant clusters word-initially and which allow for word-initial lenition, at least diachronically. In these languages, Lowenstamm's (1999) theory poses the boundary-marker at the beginning of words in the same way as in strict languages, the only difference being that the permissive boundary-marker need not be always licensed (Section 4.1). In contrast, Scheer's modification claims that the boundary-marker is not present at all in such languages (Scheer 2001, and Seigneur-Froli 2004, to appear for Greek). Considering cross-word phenomena, it must be remarked, and has been hinted at, that even in permissive languages there are rules bounded by the prosodic constituents. Recall that in the most straightforward cases, i.e., in Afro-Asiatic and Slavic languages, it is very difficult to find examples of lenition phenomena, therefore, we are forced to rely on Greek for illustrations. Nespor and Vogel (1986: 213ff) analyses Greek s-voicing as applying within the intonational phrase. Accordingly, the underlined /s/'s in (38a) are voiced, being contained within the intonational phrase (I), whereas the ones in (38b) remain voiceless since they appear at I boundaries (data from Nespor and Vogel 1986: 214).

(38) a. [$_I$ o pétros ðen íne maθiménos na trói axinús me psomí]$_I$

'Petros is not used to eating sea urchins with bread.'

b. [$_I$ ekínos o ánðras]$_I$ [$_I$ mártis mu o θeós]$_I$ [$_I$ ðen θa bi poté sto spíti mu]$_I$

'This man, God be my witness, will never enter my house.'

[39] That is why syllables and feet are bracketed in (9).

Also, nasal assimilation and stop voicing apply word-internally (39a) and across certain, but not all, word boundaries in (Demotic) Greek. For instance, they may apply between an article and a noun (39b) or the negative element and a verb (39c), but may fail to apply between an auxiliary and a verb (39d) (Nespor and Vogel 1986: 35-6, 158-159).

(39)	a. sin+pléko	si[mb]léko	'(I) knit'
	b. ton#patéra	to[mb]atéra	'the father (acc.)'
	c. ðen#pirázi	ðe[mb]irázi	'(it) doesn't matter'
	d. éxun#pléksi	*éxu[mb]léksi	'(they) have knitted'

As Nespor and Vogel suggest, these two phonological rules apply obligatorily within the phonological word, but optionally across words within the clitic group. Finally, Nespor (1986) shows that nasal plus continuant consonant sequences, evidenced in Katharevousa Greek, are rare and often lose the nasal in Dimotikí Greek word-internally (40a), across a clitic-word boundary (40b), but not across word-word boundaries (40c) (data from Nespor 1986: 67).

(40) a. anθrɔpɔs -> aθrɔpɔs 'person'

b. ðɛn θɛlɔ -> ðɛ θɛlɔ 'I don't want'

c. ɛxun ði 'they have seen'

She concludes that the p-word in Greek includes the morphosyntactic word plus its clitics. In sum, Greek seems to abound in phonological rules whose domain is smaller than the utterance. Then, if the suggestions made here are accepted, and it is the boundary-marker that blocks rule application, certain occurrences of the boundary-marker are justified in permissive languages, too. This leads us to accept that the boundary-marker is always present, as Lowenstamm says, even in permissive languages, and whether it is considered or ignored during the phonological computation is at the rules' discretion. Recall the discussion in Section 4.5.3, claiming that phonotactics

behaves in the same way as any other rule: it may choose to consider or ignore the boundary-marker. Apparently, in permissive languages phonotactics chooses to ignore it, as if it was not even there, as Scheer says. Therefore, the conclusion looks like a theoretical compromise between Lowenstamm's and Scheer's views. A closer inspection of the predictions of this analysis of permissive languages is nevertheless beyond the scope of the present thesis.

A related issue is the additional factor influencing the conditions on lenition: stress. English, and as a matter of fact Germanic languages in general, are well-known for being stress-sensitive systems, i.e., making a distinction between the onsets of stressed and unstressed vowels in their propensity to lenite, the stressed position being as strong as the word-initial one. Other languages, like Romance (e.g. French), are stress-insensitive. It is clear that the stress-sensitive vs. insensitive distinction divides strict languages into two well-defined classes, but its relation to the strict vs. permissive dichotomy is a subject for further study. The parameter responsible for stress-sensitivity was already mentioned in Section 2.1.3 (Szigetvári (1999b)'s Antipenetration Constraint), and will be elaborated on in the next chapter, where lenition in English is analysed in detail. Strangely enough, "Scheer's parameter" (see Section 4.5.5, and (34c)), i.e., the dispreference of the lenition of word-initial consonants, seems to coincide with stress-sensitivity, at least as far as the (rather limited) corpus examined here is concerned: both send Germanic and Romance into separate classes. Whether this is accidental is a question that I leave open.

A final general remark is in order here. This chapter has made an attempt at providing a unified account of language typology concerning word edges, lenition, and sandhi phenomena. Whether it is theoretically desirable at all to use the same formal devices to handle all these disparate domains of phonology is for future reconsideration. Definitely, if the attempt eventually proves to be unfruitful, this may be (partly) due to its ultimate unnecessity.

Chapter 5: Strict CV phonology and the English cross-word puzzle: the analysis

5.1 Introduction

This chapter presents a kind of climax to the discussion up to this point. It takes the theoretical background introduced in Chapters 1 and 2 as a starting point, and analyses the English data enumerated in Chapter 3, making crucial use of the modifications of Chapter 4. The key data to be considered are repeated here for convenience from Chapter 3.

(1)

a.	áṭom at^{h}ómic	Word-internally flapping only applies when the /t/ is followed by an <u>unstressed</u> vowel.
b.	t^{h}én t^{h}omáto at^{h}ómic	Both word-initial and word-internal foot-initial positions are "strong" positions, in which <u>aspiration</u> is attested.
c.	hit^{ʔ}, but^{ʔ}ler hit^{ʔ} me	The glottalised allophone appears before "a consonant and a pause", i.e. in <u>coda</u> position.
d.	hiṭ Ánn hiṭ Aníta waiṭ a mínute aṭ íssue	Across a word boundary the stress-sensitivity of flapping disappears: it also takes place before <u>stressed</u> vowels. Crucially, aspiration does not take place even in that case.
e.	a t^{h}íssue grow t^{h}omátoes	<u>Word-initial aspiration</u> is unchanged in connected speech.
f.	t^{h}o tell the truth t^{h}omorrow[1]	The beginning of <u>grammatical function words</u> is only strong utterance-initially.
g.	I want you ṭo help me Don’t lie ṭo me see you ṭomorrow	Utterance-internally lenition affects <u>t</u>'s at the beginning of function words as if there was no intervening boundary.

[1] Adverbs sometimes behave like function words, cf. (2g).

h. Please waiṭ. I'll be right back.
*They didn't waiṭ. I'll be right back.
Flapping may apply across two words in different sentences but not across just any pair of sentences. Where the two sentences are unrelated, flapping is ruled out.

Recall that in GA the most frequent environment for t-flapping (or tapping), whereby a t (or a d) turns into the voiced alveolar tap/flap [ɾ], in foot-internal intervocalic position. Thus, the t is flapped in *átom* but not in *atómic* (1a). Crucially, in the latter example the t appears in foot-initial position and gets aspirated rather than flapped. T's occurring in what is traditionally analysed as syllable-final position, i.e. in the "notorious" _{C, #} environment, show up in the form of a third allophone, glottalised [tˀ] (1c). In connected speech word-final t's seem to "ambisyllabify" into the following vowel-initial word since in a sequence such as *wait a minute* the underlined t undergoes flapping instead of glottalisation (1d). A surprising fact is that flapping loses its sensitivity to stress beyond the word level, thus word-final t's flap across-the-board within the phonological utterance while all word-initial t's remain aspirated: cf. *a* [tʰ]*issue* but *a*[ɾ] *issue* (1d-e).

In addition, t's exhibit some variation when they participate in the formation of consonant clusters: they are glottalized in *atlas*, aspirated in *chapter*, in both *attráct* and *pétrol*, tapped in *party*, and plain unaspirated in *custard*. In *winter* and *shelter*, the /t/ is tapped if the preceding sonorant is vocalized or else aspirated.

Most of the previous analyses sketched out in Section 3.2 crucially depend upon the notion of *rule ordering*. The basic question which I address here is whether a non-derivational theory is just as able to tackle the problem, revealing what the status of t in *a tissue* and *at issue* differs in, as well as what the t's in *atom* and *at issue* share. The starting point that I take draws on Rubach (1996), who observes that several connected speech phenomena can be accounted for more insightfully with reference to ambisyllabicity rather than resyllabification. In Section 5.2 I present his examples from RP, and show how GA t-tapping fits into the picture; then, in Section 5.3 I provide a Strict CV phonological analysis in the spirit of the treatment of the boundary-marker in Chapter 4. As already mentioned several times previously, I suggest that licensing and

government should be restricted to the CV-tier and the segmental level/nuclear projection, respectively. As a side-effect, the discussion reveals additional evidence against a VC skeleton (cf. Section 2.2). Section 5.4 takes a slight detour and brings supporting arguments for the analysis from the allophony of /l/ in the dialects of English. Finally, Section 5.5 concludes the chapter.

5.2 Ambisyllabicity rules in English

Ambisyllabicity is a term from traditional syllabic theory denoting a situation when a segment simultaneously belongs to two syllables, namely, when a consonant acts both as a coda and an onset. Several arguments have been put forward for ambisyllabic consonants in English, since certain segmental alternations (e.g t-aspiration/ glottalisation/tapping, r-deletion/tapping) require a three-way distinction between onset vs. coda vs. ambisyllabic consonant (see, e.g., Kahn 1976, Rubach 1996, Section 3.2.1, and the discussion below). The question that this section is considering is how Strict CV phonology, a syllabic theory stripped of branching constituents and the separation of timing and constituency (and also, by the way, of the syllable as such), is able to account for the behaviour of consonants traditionally labelled 'ambisyllabic'.

As we have seen in Section 3.2.1, Rubach (1996) refers to Kahn's (1976) study of GA, claiming that flapping affects ambisyllabic t's, and adds further data from RP (ibid: 220), which also support that ambisyllabicity rather than resyllabification applies across word boundaries in English. His examples of RP r-tapping are repeated in (2) from Chapter 4 for convenience.[2]

[2] The same data are described in Jones (1960), and Wells (1992: 282) as "a useful diagnostic" of U(pper-crust) RP as opposed to mainstream RP.

(2) RP r-tapping

a. [ɾ]: *courage, very, sorry, baron, laurel*

b. [r]: *courageous, reduce, red, bright, Henry, walrus*

c. [ɾ]: *for example, for instance, the other end*

Notice that if word-final r's were resyllabified into the following vowel-initial word, then they would get into onset position and be expected to exhibit the approximant allophone, as in (2b), and definitely not the tap. However, what we observe is the opposite (cf. (2c)), which leads us to suppose that the r's appear in the same environment in (2a) and (c). Note also that while in (2a-b) the place of stress *does* make a difference (compare *cóurage* and *courágeous*), in (2c) we get the same allophone irrespective of whether the capturing vowel is stressed (as in *for ínstance*) or not (*for exámple*). All these observations must evoke a feeling of déjà vu because of the discussion of flapping above. In (3) an additional set of data is found concerning another phenomenon in RP, l-darkening (also described in Rubach 1996).

(3) RP l-darkening[3]

a. alveolar ("clear") l: *leap, ballad, delicious, killer; kill Ann, kill a rat*

b. velarized ("dark") l: *fall, fault, marble, Marble Arch, kill, kill Joe*

Again, l's appear dark in exclusive rhyme (including nuclear and coda[4]) position, (3b), and clear both when exclusively in the onset and when ambisyllabic, (3a). Since l does not have separate allophones for pretonic and non-pretonic environments, its variation cannot serve as an argument for (or against) ambisyllabicity (as opposed to resyllabification). For the sake of symmetry, however, we will suppose that the l's in *kill Ann* or *kill a rat* are not resyllabified but ambisyllabic, thus falling under the same rubric as flapped t's and tapped r's. Thus, we can conclude that word-final consonants seem to be ambisyllabic due to the effect of a following vowel-initial word, and

[3] Below, Section 5.4 is devoted to l-darkening in the dialects of English.

[4] Syllabic l's are supposed to reside in nuclear position. See Section 5.3.4 for a discussion.

although ambisyllabification may only take place within words if the vowel after the target consonant is unstressed, there is no such constraint on the stressedness of surrounding vowels at word boundaries (cf. Kahn 1976).

Recall from Section 3.2.4 that in CM+ consonants are in a strong phonological position, which means aspiration for a /t/, when they are licensed but ungoverned; this situation emerges before stressed vowels (owing to the Antipenetration Constraint) and word-initially (when the vowel's governing potential is used up to silence the empty v in the boundary-marker). There are two types of phonologically weak positions, one is before an empty v, which is roughly before a consonant and word-finally[5] (recall _ {C, #}) – in such cases consonants remain ungoverned and unlicensed and exhibit consonantal lenition, i.e. t's are glottalised. The other weak position is that of foot-internal intervocalic C's, which receive both government and licensing from the following (unstressed) vowel; here consonants tend to move towards vocalicness, e.g. GA t's are flapped.

The question is how and why the conditions of word-final (= ungoverned unlicensed) C's change when followed by a vowel-initial word in connected speech. Recall that we have identified the position consonants take there as the same ambisyllabic context as that of foot-internal intervocalic C's on the basis of their choice of allophones (flapped t, tapped r, clear l[6]). If, for CM+, ambisyllabicity means simultaneous government and licensing, then it follows that these 'linked' consonants must also be both governed and licensed.

However, CM+ (accompanied by VC Phonology) here clashes into its own restrictions on the government capacity of vowels, and predicts that even across word boundaries there should be a difference in behaviour between stressed and unstressed vowels. When the following word starts with an unstressed vowel, it is able to exert both its government and licensing capacities upon the word-final /t/, causing it to lenite vocalically, i.e., tap (5a). Obviously, if stress domains begin with the head vowels themselves, as claimed in VC phonology, then the Antipenetration Constraint (see (4)

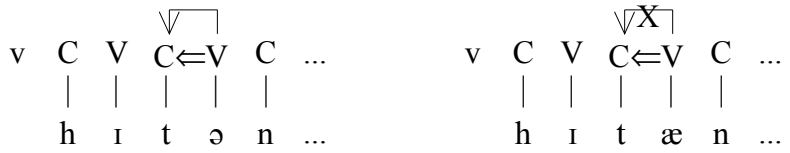
[5] In VC Phonology, word-final C's are unlicensed and ungoverned because they are followed by nothing.
[6] Although the appearance of clear l is not a straightforward indication of its being ambisyllabic (see the discussion above).

and Section 2.1.3) prevents the government emanating from a stressed vowel from affecting a preceding word-final consonant. Consequently, the /t/ in (5b) is licensed only, i.e., strong, and should be aspirated, contrary to fact.

(4) The Antipenetration Constraint

Government cannot penetrate a stress domain.

(5) a. *hit Aníta* b. *hit Ánn*

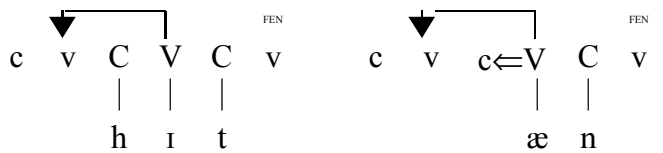


Even if we replace the VC skeleton by strictly alternating CV's, and modify the Antipenetration Constraint as is proposed in Section 2.1.3 and given as (4') below, we face serious problems. First, the government coming from the initial vowel of the second word is still diverted by the empty v of the boundary-marker required to be silent. Second, licensing cannot reach the position of t̲ either. Notice that although (6) shows *hit Ann*, the same happens in *hit Anita*: since the government of the boundary-marker does not violate any version of the Antipenetration Constraint, no difference is expected to yield from the stressed-unstressed distinction. FENs are silenced by parametric FEN-government, indicated by FEN.

(4') The Antipenetration Constraint – revised version

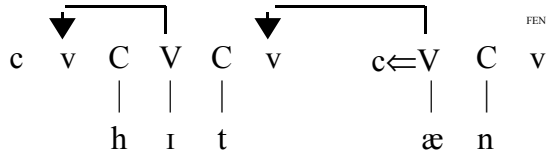
Stressed vowels are unable to govern into non-peripheral units.

(6)*hit Ann*



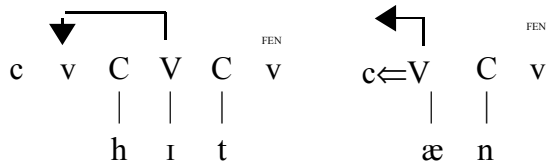
Recall from Chapter 4 that the boundary-marker is proposed here to be subject to "extraprosodicity" within the limits of the domain of application of a given rule. In addition, we saw in Section 3.2.2 that Prosodic Phonology has identified tapping as a U-level rule, i.e., one bounded by the phonological utterance only. It follows, then, that no boundary-marker is visible to /t/-allophony within the utterance[7]. Still, the closest position for the vowel's government to hit is the word-final empty nucleus, and the licensing remains trapped again (7).

(7)*hit Ann*



In Section 4.5.7 it was argued that FENs remain FEN-governed (i.e., silenced by parameter) in connected speech in English, at least before vowel-initial words. Accordingly, the vowel in *Ann* does not need to take care of the FEN closing *hit*, and the only target left for its government is the immediately preceding c position. As can be seen in (8), the word-final /t/ still stays unaffected.

(8)*hit Ann*



Therefore, we must conclude that some major modification to the model is called for, to enable it to subsume cross-word instances of lenition.

[7]Therefore it is not indicated in (7) and (8).

5.3 The analysis

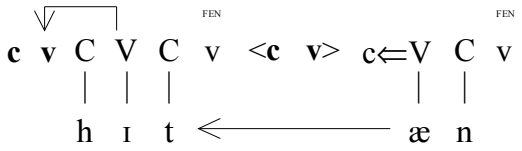
5.3.1 Ambisyllabicity again

My point of departure is the observation that ambisyllabicity across word boundaries is restricted to situations involving a vowel-initial word; word-final plosives, for example, are never reinterpreted as to their syllabic position when a liquid-initial word follows even if that plosive-liquid cluster would qualify as a possible branching onset in the language (cf. Kenstowicz 1994: 281). What this might suggest is that it is a characteristic of word-initial vowels (irrespective of whether they are stressed or not) to capture word-final consonants. I propose that cross-word left capture is motivated by the fact that initial vowels lack an onset, which results in their inability to exert their governing potential. Since it is only the *melody* of such onsets that is missing, it follows that government makes reference to adjacency on the melodic tier.

Recall from Section 4.4.2 that there is a difference in the adjacency of a consonant and a vowel when they are tautomorphemic and when a word boundary intervenes. Even if the boundary-marker is thought to be extraprosodic, a word-final consonant is still separated from a following initial vowel by a sequence of an empty v (the FEN) and an empty c. This predicts that the two situations, viz. CV and C#V, are never identical. In addition, I suggest that the government whereby a vowel affects a preceding *consonant* operates on the melodic level[8], which I base on the observation made above that V-to-C government makes reference to adjacency on the melodic tier. Notice that locality (see Section 1.2) is still observed: a word-final consonant *is* adjacent to the vowel starting the next word on the segmental level. Cf. (9).

[8] Proper Government, the other form of government, silences *vocalic* positions and is traditionally assumed to operate on the level of nuclear projection, cf Section 1.2. We have no reasons not to accept that assumption.

(9) *hit Ann*



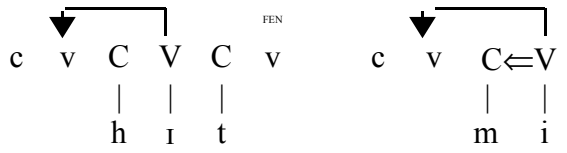
Government, then, is local, although not strictly local, since strict locality applies to adjacency on the skeleton, and thus characterizes licensing only. Notice, however, that this is not a completely novel idea: government has never been strictly local in GP anyway. PG, i.e., V-to-V government is claimed to apply on the level of nuclear projection instead of the skeleton already in Standard GP (see Section 1.2). Therefore, government and licensing do not only differ in the kind of influence they exert on neighbouring segments (i.e., damage or support), but also in the level of the structure on which they apply. Licensing is associated with the CV-tier, whereas governing relations are established either between vocalic positions adjacent on the nuclear projection, or between a V and a C adjacent on the melodic tier.[9] [10] This in turn predicts that the conditions on their application also differ. While licensing is an inherent ability of V *positions*, which may be diminished in inverse proportion to the segmental complexity of the vowel, and may even be null in the case of empty v's, certain languages may choose to treat certain empty v's, e.g. FENs, as nonempty and grant them a full licensing potential. In contrast, the governing capacity of a V position depends more on *melody*, on the internal make-up of the segment it hosts: empty v's are unable to govern, and even FENs, should they be entitled to do so, can only govern underlyingly empty v's (i.e., languages like English allow for consonant clusters word-finally) but never nonempty v's (i.e., no "deletion/syncope" is triggered by the presence of a following

[9] C-to-C government is claimed to take place between melodies by Scheer (1996 and subsequent publications): recall the suggestion that Infrasegmental Government (IG) binds consonants into a closed domain. A similar idea is pursued in Rubin (2001, to appear) in the analysis of word-final clusters.
[10] In the case of government, a similar complementariness is assumed in Charette's (1991: 91ff) analysis of *h*-aspiré between governing relations on the nuclear and the skeletal tier.

empty v)[11] or C positions (i.e., utterance-final and preconsonantal consonants are not expected to undergo vocalic lenition). In addition, I know of no case when a distinction between "better" or "worse" governors is needed, parallel to the one between licensors, defined on the basis of some aspect of the melody of the vowel. Simply, as long as there is some melody, it is able to govern.[12] Finally, the stressed/unstressed dichotomy familiar from English and expressed in the (revised) Antipenetration Constraint in (4) and (4') may also turn out to be reducible to a difference lying in the melodic identity of vowels rather than in their metrical function: unreduced vowels having no rhythmic prominence are just as unable to govern into a non-peripheral unit as their foot-head peers; witness the aspiration and lack of lenition in words like *hesit̲āte*, the impossibility of syncope in *sepa̲rāte* (verb) as opposed to *sepa̲răte* (adjective), etc. In sum, government seems to be more constrained by the melodic structure of vowels[13], which is taken here as evidence supporting the claim that V-to-C government is restricted to the melodic tier.

In (10) a situation involving a consonant-initial word is illustrated. Recall "Scheer's parameter" from Sections 4.4.1. and 4.5.5, i.e., the question whether the boundary-marker of consonant-initial words can be extraprosodic, which is set to "no" in English. Accordingly, the cv of the second word is always visible, and its empty v must be silenced by PG, so the word-initial consonant will be licensed only – therefore find itself in a phonologically strong position.

(10)*hit me*



Note that PG and the government affecting consonants are just two forms of the same relationship, as in previous analyses (i.e., CM and CM+), which means that the same

[11] Cf. Scheer (to appear).
[12] For the intriguing cross-linguistic behaviour of the schwa, see, e.g., Cyran (2003) or Scheer (2004).
[13] Scheer's IG is also a function of the melodies of the participating consonants, of course.

vowel cannot exhibit both at the same time. What follows from this is that although /m/ and /i/ are adjacent in (10) at the melodic level too, /m/ is not governed since the vowel's proper governing obligations are superior to the one hitting consonants: the ECP of the boundary-marker must be satisfied or else the whole word is ungrammatical. Since the beginning of the word is a strong position for consonants irrespective of whether the following vowel is stressed or not, (10) is claimed to be the appropriate representation for both cases. Unstressed vowels govern unconditionally, whereas stressed vowels are more restricted, at least in stress-sensitive systems like English. In VC Phonology stressed vowels are able to silence the boundary-marker because it does not reside in a separate stress domain so the Antipenetration Constraint (4) is irrelevant. Here the revised version of the Antipenetration Constraint (4') has the same result since the boundary-marker counts as a peripheral unit.

On the basis of what we have seen we can conclude the following: consonants appear strong when licensed only (cf. /m/ in (10)); lenite vocalically (e.g. t's and r's tap) when "ambisyllabic" in traditional terms, i.e. governed (/t/ in (9)); and lenite consonantally (e.g. t's are glottalised) when neither licensed nor governed (/t/ in (10)). This is more or less in accordance with CM+, although there is one essential difference: while in CM+ vocalic lenition is caused by simultaneous government and licensing, here government is enough to produce the same effect.

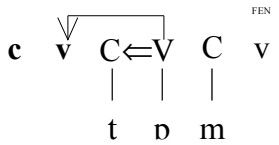
The following sections take the environments of /t/ described in Section 3.1 one by one. Section 5.3.2 repeats why the foot-initial position witnesses aspiration, and scrutinizes the Antipenetration Constraint. Section 5.3.3 provides the representation of foot-internal intervocalic positions, with the treatment of syllabic consonants adopted from Toft (2002), while Sections 5.3.4 and 5.3.5 analyse preconsonantal and postconsonantal positions, respectively. They do not present a whole new analysis but rather a synthesis of previous accounts (with ideas imported from Kahn 1976 and Harris and Kaye 1990) taylored to the version of CV phonology introduced in the foregoing chapters. Since the discussion of the representation of consonant clusters is inevitable at this point, these two sections present a tentative approach to that issue, too. They heavily rely on Scheer's (1996 and subsequent publications) and Szigetvári's (1999b)

insights. In the end, Section 5.3.6 looks into the behaviour of word-final /t/ both at the edge and within the utterance, summarizing the solution of the English cross-word puzzle.

5.3.2 Foot-initial position

Recall from Section 3.1.2 that the foot-initial position supports aspiration, both in absolute word-initial (stressed and unstressed) and word-internal pre-stress environments. The keywords are *Tom*, *tomorrow*, and *atomic* (as opposed to *atom*). As emphasized several times above, the absence of word-initial lenition stems from the fact that due to the negative setting of "Scheer's parameter" the boundary-marker is always fully present and requires silencing, cf. (11). Before stressed vowels within words, the t̲ escapes government again since the Antipenetration Constraint (4') is operative: a word-internal CV unit is non-peripheral (12b). Attested vocalic lenition (i.e. lenition by government) before unstressed vowels word-medially is also accounted for since the Antipenetration Constraint only applies to stressed vowels (12a).

(11)



(12) a. *átom* b. *atómic*

Therefore, in (12a), the word-initial vowel (/æ/) is stressed so it licenses the preceding empty c position, but since it is empty, the vowel has the potential ability to govern some other consonantal material at the melodic level (indicated by the broken single arrow), which would result in "ambisyllabicity" across word boundaries (see Section 5.3.6). The second vowel (/ə/), however, being unstressed, is able to govern the preceding consonant, and incidentally, it also happens to license it. Hence, the /t/ in *atom* will be tapped and so will the underlined /t/ in e.g. *hit atoms*. In (12b) the word-initial vowel is not stressed, thus tries to govern, which will not materialise until the word is put into such a context where it is preceded by a consonant-final word, e.g. *hit atomic elements*. In that case government reaches the underlined /t/ surfacing as a tap. At the same time, the initial empty c position consumes the vowel's licensing. The stressed vowel in (12b), on the other hand, licenses the /t/ making it aspirated, but cannot govern it (in accordance with (4')).

The degrees of aspiration owing to the stress degree of the vowels in *ten!, ten, temperamental, latex, hesitate* (see Sections 3.1.1 and 3.1.2) may result from the gradability of licensing, referred to in the previous section. The more prominent a vowel is the better licensor it becomes. Notice that, as argued above, this difference in licensing capacity is not connected to the melodic make-up of the vowels but the degree of prominence: the same vowel quality is emphatically or non-emphatically stressed (cf. *ten!* and *ten*), and the same set of vowels serves all the non-zero stresses (cf. *ten, latex, hesitate*).

In connected speech, word-initial /t/ remains strong and aspirated, as in *buy tomatoes*, due to the setting of "Scheer's parameter". The beginning of function words like *to*, however, is only strong utterance-initially since function words in English normally cliticize onto the preceding word and form a single p-word with it. Consequently, the boundary-marker at the left edge of function words ceases to denote a word boundary, and it falls out of the scope of "Scheer's parameter": it is ignored by the phonology, and the vowel's government hits the immediately preceding consonant. Such function words always contain unstressed vowels (being unstressed, i.e., pronounced in

the so-called weak form, is a condition on their cliticization), and thus the Antipenetration Constraint is inapplicable. Therefore, phrases like *lie to me* fall under the rubric of foot-internal intervocalic positions, and are accordingly discussed and illustrated in the next section.

It is worth warning the reader at this point against concluding that the present discussion assigns some theoretical status to the foot. On the contrary, the existence of metrical structure in the form of hierarchically arranged constituents is denied here in the same way as in Standard GP. The term *foot*, then, is used here as a mere descriptive label. Notice that we have not once needed reference to the *whole* of a foot but its edges only: in a trochaic system like English, the left edge is the position immediately preceded by the stressed vowel, the right edge is the one followed by another left edge if followed by anything at all.

In the rest of the section the Antipenetration Constraint is under scrutiny. It has been suggested above that it has the status of a parameter: it simply seems to generate cross-linguistic differences irrespective of the other characteristics of the phonological systems. It has also been alluded to above that in Germanic languages at least, the stress-sensitivity of lenition stemming from the Antipenetration Constraint and the negative setting of "Scheer's parameter" correlate in a way not fully understood yet. In Section 5.2 we proposed a modification to the Antipenetration Constraint, replacing the original reference to stress domains with the reference to the distinction between peripheral and non-peripheral units (adopted from Szigetvári 1999b: 100). This will prove useful, as is shown below, in not only accounting for the cross-word weakening of final consonants but also in handling another configuration: words with a stressed first vowel in connected speech.

Suppose we are working with a VC-skeleton armed with the original version of the Antipenetration Constraint: government cannot penetrate a stress domain (4). In English, "feet" are trochaic, therefore stress domains start with the stressed vowels themselves and reach up to the following stressed vowel. Crucially, in VC phonology the consonants immediately preceding the stressed vowels do not belong to the same stress domain. Initial peripheral units, i.e. vC at the beginning of words, do not form a

separate stress domain and thus government is not prevented from entering them. It follows, then, that it is the right boundary of stress domains that is decisive with respect to the Antipenetration Constraint: in *at*)*ómic*[14], there is another stress domain to the left of the stressed vowel and consequently it is unable to govern into it; in v*Tóm*, there is not and consequently the stressed vowel *is* able to do so. Therefore, the /t/ in both *atomic* and *Tom* is protected from government and strong; the governing capacity of the stressed vowel in *atomic* is left unexploited, while in *Tom* it is used to silence the initial empty v. Notice that should the Antipenetration Constraint refer to the left boundaries of stress domains, no difference would exist between *at*(*ómic* and v*T*(*óm*, the empty v position at the beginning of *Tom* would be unsilenced, and its non-pronunciation would violate the ECP.

Next, consider *hit*) *Ánn* and *h*)*ít at*)*ómic elements*, where I only indicate the stresses and stress domain boundaries which are relevant to the present discussion. As shown above, the theory predicts no tapping in *hit Ann* since government cannot cross right boundaries and therefore cannot affect the word-final /t/. Even in *hit atomic*, the /t/ of *hit* will only be governed if we assume that stress domains are either restructured postlexically or, as Scheer would probably put it, are only established postsyntactically, which results in /t/ being domain-final in *hit Ann* but not in *hit atomic*. In *hit atomic*, the /t/ of *hit* is thus contained within the stress domain headed by /ɪ/ (in exactly the same way as in, say, *bitter*) and hit by the government emanating from the initial unstressed vowel of *atomic*, and undergoes tapping.

Finally, consider the situation when *Tom* is preceded by another word, e.g., *see Tom*. In light of the previous discussion, the stress domains are structured like this: v*sée*c v*T*)*óm*. That is, everything before the stressed vowel in *Tom* gets incorporated into the preceding stress domain, just as in *hit atomic*. However, while in *hit atomic* this move ensured the government of the /t/ in *hit*, in this case it erects a right boundary between the /t/ and the /ɒ/, which obstructs government. As far as the /t/ is concerned, it should in fact escape government and remain strong, but the empty v before it is also

[14] The boundaries of stress domains are indicated by parentheses, following the convention in Szigetvári (1999).

left orphan and its ECP unsatisfied. In VC phonology such situations, i.e. When an ungoverned empty v seems to be phonetically unrealized, will always violate the ECP unless some major modification or amendment is introduced into the theory.

CV phonology operating with the revised version of the Antipenetration Constraint ((4'): stressed vowels cannot govern into non-peripheral units) never faces such a problem. The boundary-marker at the beginning of cv*Tom* (underlined) is a peripheral unit and therefore governable even by stressed vowels; and so is the end of consonant-final words like *hit*v. In *atómic* the /t/ resides in a nonperipheral unit and escapes government; and finally, whenever *Tom* is used, the boundary-marker is always there (recall "Scheer's parameter") and the status of the initial /t/ does not change.

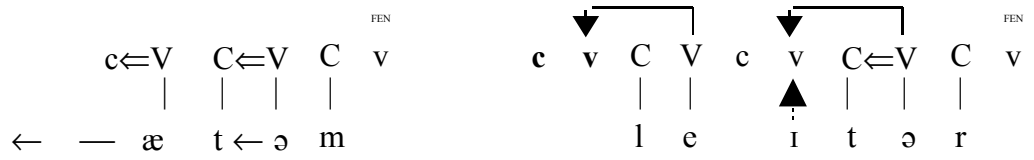
The revised Antipenetration Constraint may sound arbitrary and stipulative, which simply reflects the fact that it is really not more than the statement of an observation, and has not much explanatory power. Notice however, that the original version of the Antipenetration Constraint is not much less stipulative either; after all, why should government, rather than licensing, be incapable of penetrating stress domains; what is it about stress domains that makes them difficult to penetrate; and, last but not least, what *is* a stress domain in a flat phonological theory like CV/VC, and what is its relation to the traditional concept of "foot"? Questions like these are not much less grave than those concerning the revised version, e.g., what is the essential difference between peripheral and nonperipheral units that ultimately derives the stress sensitivity of a lenition system. Presently, I have no answers to such questions.

5.3.3 Foot-internal intervocalic positions

Section 3.1.3 discussed /t/-allophony in foot-internal intervocalic positions, and identified the following key data. The intervocalic position immediately following the (either primary or secondary) stressed vowel seems to be a relatively straightforward environment for tapping (e.g., *átom*), although some speakers have a tap after lax vowels only but not after tense vowels (e.g., *writer, later*). In (13a), the representation

of *atom* is repeated for convenience: the /t/ is governed (and, incidentally, also licensed) and therefore undergoes vocalic lenition. The difference between short lax vowels and long, diphthongal tense vowels resulting for some speakers in the difference in the tendency to tap in *atom* but not in *later* must reside in the representation of long vowels. As mentioned in Section 2.2.1, in CV (or VC) phonology long vowels are considered to have the underlying structure Vcv, where the first V is lexically occupied by the melody of the vowel, but the second vocalic position must satisfy certain structural conditions to become available for the spreading of the first V's melody (in long monophthongs) or for the docking of some floating material (in diphthongs). In the earliest versions of CV phonology (e.g., Lowenstamm 1996, Larsen 1995), that structural condition is the PG emanating from the following nonempty V (or a FEN in languages where it is parametrically enabled to (properly) govern). If we accept this[15], then it readily explains the lack of tapping in words like *later*: the government of the second vowel is used up by the first vowel instead of the consonant sandwiched inbetween, cf. (13b).

(13) a. *atom* b. *later*



In light of this, it suddenly becomes mysterious how tapping is still possible for some speakers in words like *later*. Notice, however, that the present theoretical framework has no apparatus for handling cases of variation like the one at issue: it either covers the dialects which tap in *later* (by assigning some representation to long vowels different from (13b)) or the dialects which aspirate at the same location (by accepting (13b)), but not both. The choice would have to be determined by a cross-dialectal and cross-linguistic study of intervocalic consonantal behaviour.[16]

[15] Although this view has been criticized by both Szigetvári (1999b) and Scheer (2004). See Section 2.2.1.
[16] In Scheer (2004) V-to-V *licensing* is the relation that ensures the pronunciation of long vowels. Notice that this proposal makes the opposite prediction: that the intervocalic position cannot be strong after a long vowel.

Section 3.1.3 also introduces the so-called "Withgott-effect": the non-flapping of t̲'s in examples like *milit̲aristic, Medit̲erranean, Navrat̲ilova*. These contain t̲'s in the third position of word-internal dactylic sequences, whose phonological strength has been accounted for with reference to footing (e.g., Withgott 1982, Jensen 1987, 2000, Davis 2004) and to Paradigm Uniformity (e.g., Steriade 2000). Section 3.1.3 argues for similar effects in final dactyls like *sanity*. Since this phenomenon stems from a complex interplay of metrical structure, government and licensing, its discussion is postponed to Chapter 6, which provides the supporting evidence of the *sanity-militaristic* symmetry, introduces and evaluates previous analyses, and offers an alternative.

One set of data with intervocalic /t/ that will not be analysed here contains the problematic -*to* words first discussed within the generative paradigm in SPE (pp. 190-191). It is obvious that the fact that flapping takes place in *motto* whereas aspiration is attested in *grotto* can be handled by no theory without access to arbitrary lexical marking, and therefore no attempt will be made to do so.

Following syllabic consonants, however, do cause a systematic difference: between t̲'s before a syllabic /l/ (e.g. *battle, little*)[17], where tapping *is* attested, and t̲'s before a syllabic /n/ (e.g. *button*), where glottalisation is frequent instead. The difference between the two syllabic consonants was also revealed by Toft (2002), in her phonetic study of syllabic sonorants in Southern British English. She found that /l/ is syllabic irrespective of the preceding context, while the distribution of syllabic /n/ is context-dependent, after both single consonants (e.g. *bottle* and *button*) and consonant clusters (in words like *dwindle* and *London*). With respect to duration, she argues that syllabic /l/ patterns like onset /l/, not like coda /l/, nor as a distinct category; whereas syllabic /n/ patterns like coda /n/, and not like onset /n/, nor as a distinct category. In her recordings, words with a potential syllabic /l/ never contained a schwa and the /l/ was syllabic in all the examples, while words with a potential syllabic /n/ vacillated as a function of the context. As she advocates Standard GP, she concludes that the representations of the two syllabic sonorants are as shown in (14), i.e., syllabic /l/ is exclusively attached to a

[17] Syllabic /m/ (as in *bottom*) and /r/ (as in *butter*) behave analogously.

nuclear position whereas syllabic /n/ is generated in an onset but spreads to the left to satisfy the ECP of that empty nucleus.

(14) a. b.

```
N        N        O
|        |        |
x        x        x
|         \       |
l          \      n
```

Suppose that Toft's findings also apply to the tapping dialects in question. Then it is correctly predicted that tapping more readily takes place before a syllabic /l/, since it occupies a vocalic position, therefore that situation simply reduces to the intervocalic case. That is, *battle* (15a) is expected to exhibit the same behaviour as *atom* does (cf. (12a)). In contrast, whenever there is a potentially syllabic /n/, e.g., *button*, an empty v intervenes between the /t/ and the /n/, whose ECP is satisfied either by phonetic realization as a schwa, and then tapping may ensue; or by the spreading of the /n/ (15b). It appears that sharing the melody of the nasal with the following C position is not enough for the empty v position to acquire the potentials of a nonempty V: it is still unable to govern or license, and as a consequence, the preceding /t/ surfaces as glottalized rather than tapped.

(15) a. *battle* b. *button*

```
                                                FEN
C   V   C⇐V          C   V   C   v   C   v
|   |   |   |        |   |   |   \   |
b   æ   t ← l        b   ʌ   t       n
```

Finally, since unstressed function words in English cliticize into the p-word of their hosts preceding them, "Scheer's parameter" does not apply to their boundary-marker, which is consequently extraprosodic for the tapping rule. For instance, the initial /t/ in *to* is only aspirated when at the beginning of the utterance, otherwise it is flapped if it is

preceded by a vowel-final word (e.g., *lie to me*). (16) illustrates that ... *lie to* ... creates exactly the same context for the /t/ as *atom* does.[18]

(16) a. *atom* b. *lie to*

```
                                FEN
      c⇐V  C⇐V  C  v              ...  V  c  V  C⇐V
        |    |   |  |                   |     |   |   |
←  —  æ   t ← ə  m                       a     ɪ   t ← ə
```

Such t-initial function words illuminate a significant difference between CV and VC phonology. Even if the extraprosodicity account should prove to be incorrect in due course, there are just two ways of explaining how the identity of the left edge of function words changes from strong (utterance-initially) to weak (utterance-medially), or vice versa. Either all words including functional elements carry the empty CV span, which is deleted within the utterance; or pre-function word boundary markers are absent, and the empty CV unit is inserted utterance-initially. What seems to be clear is that a VC analysis fails in either case. Since in consonant-initial words it is the V position of the first VC unit that functions as a boundary marker (i.e. it absorbs the governing potential of the following nonempty V), it can be neither deleted nor inserted under any circumstances as "VC units [...] are here claimed to be inseparable" (Szigetvári 1999b: 108).

5.3.4 Preconsonantal positions

The representation of the different types of consonant clusters[19] has always puzzled CV/VC phonologists. This is because, in a strict CV/VC skeleton, any two C positions potentially forming a relationship are separated by an empty v. Therefore, all consonant clusters are "bogus" in the Standard GP sense of the word, i.e., the members are never

[18] The extraprosodic boundary-marker is ignored for the sake of simplicity.
[19] In the forthcoming discussion, traditional labels like "onset" and "coda", "onset cluster" and "coda-onset cluster" are used in a merely descriptive sense.

strictly adjacent (see Chapter 1). Consequently, the need for new mechanisms silencing the intervening v's arises. Besides PG, already in Standard GP the notion of O-to-O government is thought to apply in certain CvC sequences, having the power to satisfy the ECP of the empty nucleus sandwiched inbetween, and we have also referred to Scheer's (1996) Infrasegmental Government (IG), which produces a closed domain with the same result.[20]

There are a few issues, though, which divide even CV/VC-ers, and largely influence the way they conceive of consonant clusters. Firstly, it is not at all clear how many types of cluster are to be distinguished representationally. While Szigetvári (1999b) insists that "based on their behaviour at least three different types of consonant cluster must be distinguished" (ibid: 111), which is basically a reflection of the traditional GP classification into branching onsets, coda-onset interconstituent domains, and bogus (fake) clusters, later (Szigetvári 2002a-b, to appear[21]) he makes an attempt at reducing the number to two by subsuming onset clusters and (syncope-created) bogus clusters under the same heading. Meanwhile, Scheer (e.g., 2004) firmly claims that the number is indeed two, although it is coda-onset clusters and bogus clusters that he does not distinguish: both are simple CvC sequences with PG silencing the empty v and no special relationship contracted by the consonants themselves, as opposed to onset clusters, whose terms enter into the IG frequently referred to above.

A second debatable issue concerns the direction of the relationships consonants may form, and the variability found in Standard GP (constituent-government is head-initial, interconstituent government is head-final) is replaced by fixed directionality, which is right-to-left (cf., e.g., Scheer 1998b). Once the direction of consonantal interaction is unified, however, the analyst is forced to choose between Standard GP's head-initial and head-final domains, since both cannot be accounted for at the same time due to the well-known observation that they are each other's near mirror images. Since in GP consonantal interaction has always been a function of the elemental make-up of the participants (cf. Charm Theory or the Complexity Condition in Section 1.5), one

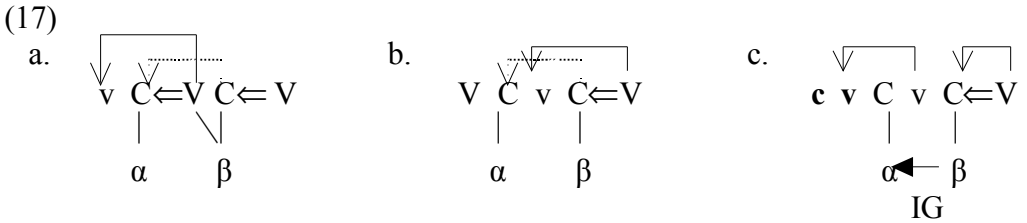
[20] For the special status of IG, see Scheer (2004: 162-163).
[21] In fact, as early as page 120 of Szigetvári (1999b), he already hints at the possibility that "word-internal onset clusters are a special type of bogus cluster".

either reverses the complexity relations of the segments of onset clusters in a revised theory of elements, as Scheer does, and then interconsonantal government derives those (but no other) clusters. Alternatively, one may introduce *two* right-headed relations, as Szigetvári (1999b) does, one characterising onsets, the other characterising coda-onset sequences.

Although the discussion, and the solution of the problem posed by, consonant clusters in CV phonology is beyond the scope of the present thesis, the analysis of /t/-allophones requires sketching out a model which is capable of covering the data given in Chapter 3. Since that model of onset (or TR-) clusters (to be presented in (17) below) draws on both Szigetvári's (1999b) and Scheer's (2004, to appear), first these two are introduced briefly.

Consider Szigetvári's representation of word-initial and word-medial onset clusters in (17a) and (17b), respectively. The two are not identical as Szigetvári realizes that while the empty v of a boundary-marker requires PG in the initial situation, word-internal vowels must not be attacked by it since neither syncope nor bogus clusters are attested before onset clusters. In addition, he claims that the vocalic position between the terms of the cluster (α and β) must be nonempty (due to the second melody – β – spreading onto it) and able to govern the boundary-marker initially, but it is empty because it is governed medially. Furthermore, the two consonants contract a C-to-C licensing relationship (indicated by the dotted arrow).

(17)

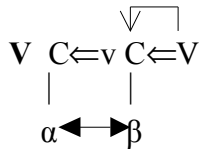


(17c) shows Scheer's representation of onset clusters. Notice that this is a slightly modified version compared to the original closed-domain analysis (Scheer 1996): it is an improvement in that it observes the locality condition on government. As Scheer

claims, empty v's enclosed within an IG relation are ungoverned and therefore endowed with full lateral actorship, i.e., it is them that govern the boundary-marker rather than the nonempty V following the cluster. Unfortunately Scheer is not explicit on how word-medial onset clusters should be represented.

Our model will not radically differ from either of the two representations; in fact, it only departs from them in order to fit the discussion of the preceding sections, namely, to further exploit the revised Antipenetration Constraint. From Szigetvári, I adopt the idea of the difference between word-initial and internal cases, although the two do not deviate to such an extent here; from Scheer, I adopt the proposal that the enclosed v is very much like a nonempty V. I subscribe to the claim that the formation of onset clusters is heavily constrained by the melodic structure of the participating consonants, which suggests that some kind of interaction exists between them, whether it is called C-to-C licensing or IG. It is denoted by the two-headed arrow in (18).

(18)



It is clear that the empty v sandwiched between the two C's *can* govern the preceding vocalic position (given in bold) if it is part of the boundary-marker, i.e., if it is in a peripheral unit, but cannot govern it if it is a word-internal position. This is reminiscent of the revised Antipenetration Constraint (4') and the characterization of stressed vowels in English: they are unable to govern into nonperipheral units. Therefore, I propose that the consonantal relation between α and β (for the sake of simplicity, let's refer to it as IG) indeed assigns special properties to the intervening v position: it behaves like a stressed vowel in conforming to the revised Antipenetration Constraint. The difference between word-initial and word-internal onset clusters, then, lies in the limited capacities of the sandwiched v: it is only able to silence a boundary-marker, but, crucially,

nonperipheral (empty or nonempty) positions are left unaffected by it, and this predicts the absence of phonetically unrealized empty v's (either created by syncope or lexically specified as empty) to the left of onset clusters. Bear in mind, though, that the claim here is that such empty v's only *resemble* stressed vowels in having the same potentials – of course, they are not stressed vowels, that is, they do not play such a role in metrical structure, they cannot since they are empty.

A look at (18) immediately reveals a difference in phonological strength between the two members of the cluster: while the first position is strong, being licensed only (the government of the enclosed v is either used up by a boundary-marker or is left unexploited due to the revised Antipenetration Constraint), the second is weak, being licensed *and* governed at the same time. This difference is desirable: recall the old assumption made by Standard GP, that government proceeds from left to right in branching onsets, which produces a distributional asymmetry between the two terms: the recessive second position allows for a limited set of contrasts. Moreover, a (licensed and) governed weak position predicts vocalic lenition in the present framework, e.g. voicing or sonorization, which is supported by the tendency for sonorants to occupy the second position of onset clusters. The first C, however, is strong, which means it is aspirated whenever taken by a voiceless plosive.

A point that bears high relevance to the discussion of aspiration in English and was discussed in some detail in Section 3.1.4 concerns the fact that most descriptions of the phenomenon find no aspiration in single foot-internal intervocalic plosives (as in *petal*), but they identify devoiced sonorants in the corresponding onset clusters (as in *petrol*), a sign frequently associated with the aspiration of the preceding plosive. Although this point is still subject to debate (cf. the discussion in Section 3.1.4), the present analysis supports the view that the /t/ is indeed aspirated in both *betráy* and *pétrol*, irrespective of stress relations: in (18), the status of the first consonant cannot be influenced by the stressedness of the vowel following the cluster, it is strong in any case. In fact, the element that may be influenced by that vowel is the second consonant: it is governed and licensed only if it is followed by an unstressed vowel (i.e., in *petrol*); if a stressed vowel follows, then the revised Antipenetration Constraint prevents it from

governing. That is, the second C position is strong when it precedes a stressed vowel (i.e., in *betray*). Whether this prediction is borne out by any (phonetic) detail of the data is an open question.

Preconsonantal t's other than the ones in TR-clusters are generally glottalized, cf. *Atkins, chutney, atlas*, etc. This is due to the fact that the consonant clusters in which the /t/ participates there are bogus (in both Szigetvári's and Scheer's models), that is, the enclosed v in CvC is silenced by the PG emanating from the following vowel, and is unable to license or govern (see (19b) below). As a consequence, the /t/ (α in (19b)) occupies a hopelessly weak phonological position: it is unlicensed and undergoes consonantal lenition, i.e., glottalisation. The same applies to syncope-created bogus clusters like *batt'ry* and *pott'ry*.

We can conclude, then, that preconsonantal /t/ is either glottalized, or it forms an IG domain with the following segment and causes it to devoice. In most cases the choice is categorical: *chutney* is only reported with a glottalized /t/, *petrol* only comes with a devoiced /r/[22]. In a few cases, e.g. *atlas* (Hoard 1971 and Section 3.1.4) and syncopated *battery* (Harris and Kaye 1990: 270 and Section 3.1.4 fn. 13), free variation is possible between the two candidate pronunciations, which suggests that the parsing of segment sequences is not always straightforward[23]. Moreover, if sonorant devoicing is considered to take place when /t/ enters into IG with the following sonorant, it is not only the /t/ of *battery* that can form that relation after syncope, but the /t/ of *atlas*, too, at least in one possible pronunciation, which is a surprising conclusion[24]. Notice that the apparent counterexamples Spencer (1996: 212-213) brings against claiming that sonorant devoicing results from aspiration are all fricative-approximant sequences that

[22] Recall from Section 3.1.4 fn. 13, that certain British English speakers display both in words like *petrol*. Also, the same section points out that the devoicing of a following /r/ is not necessarily a consequence of aspiration but may also be due to the affrication process /tr/ and /dr/ clusters uniformly undergo. However, since the other voiceless plosives devoice following sonorants in much the same way (cf. *apron, acrid*), I assume that the /t/ also has a devoicing effect on the /r/ irrespective of affrication.

[23] A similar indeterminacy is observed in the case of *sC* clusters in stress assignment: in words like *mínister, sínister*, they behave like ordinary onset clusters do, resulting in antepenultimate stress (cf. *álgebra*), whereas in *seméster*, the same cluster is analysed as a coda-onset sequence, which attracts stress onto the penult (cf. *agénda*). English is not alone in exhibiting "word-internal s+C effects", see, e.g. Scheer's (2004: 443) examples from Czech.

[24] This may again support Szigetvári's idea referred to above, that onset clusters and bogus clusters are not two distinct categories.

are analysable as traditional onset clusters, i.e., ones created by IG (*Islip, eye-slip, midwifery* (with syncope-created /fr/), *fried, shrimp, mushroom*). Crucially, no devoicing is found in *ice-lip*, where the /sl/ sequence is either a coda-onset cluster or bogus (depending on theoretical taste) but definitely not one involving IG.

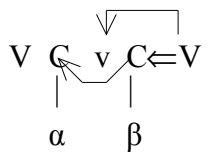
In sum, the second position in an onset cluster seems to be subject to devoicing. Why this should be the case is not clear, and I do not see which element in the representations in (17) or (18) above may be responsible for it. Yet, as is shown in the next section, it may be considered to be part of a more general phenomenon of the sharing of laryngeal features under IG.

5.3.5 Postconsonantal positions

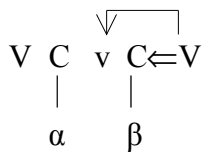
(19a) and (19b) show Szigetvári's and Scheer's representations of coda-onset clusters, respectively. Since Szigetvári draws a distinction between coda-onset and bogus clusters, (19b) also serves as the illustration of the latter type for him. Scheer ignores that distinction and uses (19b) for both. (19a) demonstrates what Szigetvári calls C-to-C government[25].

(19)

a.



b.



The notion of C-to-C government may be in conflict with the interpretation that government receives in the present theoretical framework: it usually causes vocalic lenition, which it does not in cases of (19a). This is because the theory has been relaxed

[25] One of the effects of C-to-C government is the silencing of the enclosed v, which also receives the PG of the following nonempty V if there is one. Even if there is not, the cluster is well-formed since PG is not necessary for the satisfaction of the ECP. This ensures that such (traditionally, coda-onset) clusters, but not bogus clusters (19b), can occur word-finally.

here and government alone is thought to be responsible for vocalic lenition, as opposed to Szigetvári's system, where government must be accompanied by licensing for such a result. The present analysis is thus unable to interpret C-to-C government. Nevertheless, since the account of /t/-allophony does require the distinction between (19a) and (19b) as well as recourse to a right-to-left consonantal interaction in some clusters, Szigetvári's version will be adopted this time, and the issue of how C-to-C government fits into the rest of the analysis will be left open.

Word-internally the distinction between the two non-onset clusters (19a-b) may not manifest itself in /t/-lenition. This is because whenever /t/ is the second member (*party, winter, captain*, etc.), the cluster is a coda-onset sequence where the /t/ is the onset and is strong (true for both (19a-b)); whereas whenever the /t/ is the first member (*Atkins, atlas*, etc.), the cluster is bogus where the /t/ is the "coda" and undergoes lenition (also true for both (19a-b)).

The word-final and cross-word situations, however, will unambiguously opt for (19a). To account for the absence of glottalisation in word-final position after an obstruent (*fist, fact, apt*, etc.), there are two possibilities. One either claims that the final V in (19b), the FEN in the words in question, is able to license the /t/ and that is why it does not lenite. Then it is not clear why the FEN in *hit, hurt, hint* is unable to do so. Alternatively, and this is the thread we will follow here, one accepts (19a) and claims that the /t/ has governing obligations in word-final coda-onset clusters, i.e., it forms the C-to-C governing domain with the preceding consonant. In order to do so, it must be melodically strong (or complex, as Harris 1990 would put it), which can only be attained if the /t/ remains an unlenited /t/. In cross-word cases like *fact is*, *kept it*, even the government emanating from the initial vowel of the following word is there to attack the /t/ but it cannot affect it for the same reason. This explanation is in line with the one Harris and Kaye (1990) propose for the treatment of the "protected environments" of /t/ (cf. Section 3.2.3).

Similarly, in words like *chapter* and *doctor*, the /t/ is second in a C-to-C governing domain, and as such is shielded from lenition; moreover, since it is followed

by a nonempty vowel, it is licensed this time and therefore strong and willing to aspirate.

As far as the other consonants are concerned that may precede /t/, viz. the sonorant consonants /r/, /l/, /n/, they create exactly the same environment as a vowel does: recall from Section 3.1.5 that the /t/ is glottalized in *hurt, hint* and *belt* (as in *hit*), but tapped both in *forty, winter, shelter* (as in *atom*) and in *hurt it, paint it, belt up* (as in *hit it*). Although aspiration is another option with /l/ and /n/, as also described in Section 3.1.5, Kahn (1976) finds that this happens when the sonorants have a consonantal articulation. Then we simply face subcases of coda-onset clusters like the ones in *chapter* and *doctor*, where the /t/ is protected from lenition. On the other hand, when the nasal itself is deleted or the /l/ is vocalized, we simply face subcases of intervocalic flapping: the sonorant consonant gets incorporated into the "nucleus", in the form of vowel nasalization in the case of /n/, in the form of a nonlow back offglide lending a diphthongal quality even to monophthongal vowels in the case of /l/[26].[27]

With /r/, the data is less ambiguous: tapping is generally reported as acceptable in /rt/. There is no difference between a consonantal articulation and a vocalic articulation: /r/ always behaves like a vowel. Based on such and similar evidence, several authors beginning with Kahn (1976) have concluded that English /r/ is a (semi) vowel.[28]

Section 3.1.5 points out that unstressed t̲-initial function words following consonant-final ones always parallel the corresponding word-internal situation, e.g., *talk to* = *doctor*. Again, the pre-sonorant position is evaluated according to whether the sonorant is pronounced with consonantal constriction. Whenever a true coda-onset-type of cluster is created, where the first consonant is final in the first word and the second consonant is the initial /t/ of the function word, the FEN of the first word intervenes. Recall from Section 4.5.7 that the connected speech behaviour of FENs is not fully understood yet, but at least in English, it seems to lose its parametric government across

[26] An interesting prediction is that speakers who do not flap following long vowels (including diphthongs) (Section 5.3.3) will not do so in /lt/ even if the /l/ is vocalized.

[27] Harris and Kaye (1990) point out the same free variation, even in words like *belt*, where London glottalling is more likely if the /l/ has become vocalized, i.e. /bɛɫt/ or /bɛwʔ/.

[28] Recent examples include Harris (1994), Csides (1998), and Dubach Green (2003).

words when a consonant-initial word follows. This is the case in question here. As a result, the underlined v in *talkvto* and *docvtor* receive the same treatment, i.e., they are properly governed by the following filled vocalic position.

Leslie's (1983) and Harris and Kaye's (1990: 271) "13-14" words (or the *-ee/-oo* cases, as Harris and Kaye call them) (i.e., words such as *thirteen, fourteen, eighteen, nineteen, canteen, frontier, settee; pontoon, cartoon, tattoo, spittoon; seventeen, guarantee*, in which t's lenite in pretonic position) can only be conceived of as a set of lexically marked items in the case of which the Antipenetration Constraint is suspended. This set mostly consists of highly frequent words such as numerals, so this kind of extension of lenition seems plausible.[29]

The behaviour of *sC*-clusters has been puzzling scholars for ages, so much so that Scheer (2004: 444) dubs it a "phonological Nobel-prize problem". Space limitations prevent me from summarizing all the previous suggestions – if such a summary is possible at all, considering the vastness of the relevant literature. Basically, what sets them apart from other clusters is their apparent distributional liberty in a number of languages including English: they can occur initially, medially, and finally, too. In this respect, they imitate single segments, so no wonder one possible analysis proposes that they are a kind of contour segment (e.g. Selkirk 1982, Wiese 1996). The difficulty that descriptions craving to analyze them as clusters continually run into stems from the fact that due to their distributional properties no single, uniform representation can be established. If they are analyzed as coda-onset clusters, then their word-initial occurrences require special treatment, cf. Kaye's (1992) Magic Licensing, or the discussion at the beginning of Section 4.1, written in the same spirit. If, on the other hand, they are analyzed as onset clusters, then the word-final position is problematic: why can other onset clusters like /tr/ or /pl/ not appear there in the same language? Therefore I conclude that a severely constrained phonological theory like GP, denying arbitrary operations that reorganize structure and insisting on the strong connection

[29] The connection between word frequency and the likelihood of consonant lenition and vowel reduction has been long observed (see, e.g., Hooper 1976).

between the melodic structure of consonantal segments and the clusters they form, is unable to tackle such a problem.

The theory can be relaxed, though. It is possible to claim that *sC* sequences are essentially coda-onset clusters, with C-to-C government established between the members. Their natural "habitat" is the medial and final positions. Languages, however, may choose to reanalyze such C-to-C government domains as IG domains, and allow for their word-initial appearance.[30] At this point the consonants are free to choose the most "suitable" government relation, IG word-initially but C-to-C government word-finally, in order to survive, i.e., yield a grammatical configuration. Since word-internally both are possible, the clusters will vacillate between the two constructions, cf. footnote 23. To some extent, the choice is governed by stress relations, for reasons still unclear: the majority of intervocalic *sC* clusters follow the TR-pattern, except when they are trapped between stresses, e.g., in Davidsen-Nielsen's (1974) and Jensen's (1987) examples, when there is secondary stress on the preceding syllable, e.g., *gèstátion*. In such cases, the clusters retain C-to-C government and the /t/ is free to aspirate.

This reanalysis is similar to the establishment of IG relations that accompanies the lexicalization of syncope-created bogus clusters. Recall that one possible pronunciation of *batt'ry* comes with a devoiced /r/; similarly, *s'ppose* is well-established with an unaspirated /p/, and *Wisconsin* contains an unaspirated /k/ in its Wisconsin pronunciation. Notice that the possibility of reanalysis is closely related to word frequency: *s'ppose* is definitely more frequent than *S'pir* (for *Sapir*) (at least in non-linguists' speech), reference to *Wisconsin* is undoubtedly more frequent in Wisconsin than outside that state.

The remaining question is when exactly this reanalysis can take place. Why is it by far the most frequent with /s/ or /ʃ/ as the first term? The answer must be hidden in the elemental make-up of sibilants, as Kaye (1992) also admits. Certainly, the melodic content of consonants seems to be of utmost significance in the phonotactics of all

[30] This claim, that languages may decide to analyze consonant sequences in one way or another in a more or less arbitrary fashion, is also supported by data from the loan adaptation of clusters. As Fleischhacker (2001) has shown, languages exhibiting an anaptyxis-prothesis asymmetry may divide consonant clusters into the two patterns at various points of a continuum of cluster types, which is (roughly) organized according to a traditional sonority hierarchy.

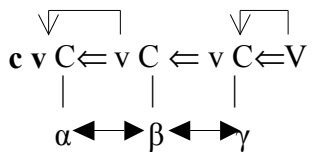
clusters. It cannot be an accident that *sC*, the exceptional initial cluster, is one of those non-rising-sonority clusters which can unambiguously be analyzed as coda-onset, rather than one of those we categorize as bogus.[31] This lends additional support to the reanalysis account: it is easier to reanalyze an existing relationship than to establish a new one.

The reason why there is no aspiration in *sC* may lie in the long-observed phonetic fact that such clusters are produced with a single glottal movement (cf., e.g., Kim 1970 and Section 3.1.5)[32]. This is reminiscent of the sharing of laryngeal features under IG as in /tr/, /pl/, etc., proposed in the previous section. If we claim, as is done above, that under certain circumstances *sC* strings may be "translated" from C-to-C government domains into IG domains, then it is not surprising any more that aspiration is not inhibited whenever IG cannot be contracted, e.g., in the case of morphologically transparent ("analytic") s-final prefixes as in *miscalculate*. Recall from Section 4.5.3 that the domain of application of phonotactics is the word: the boundary-marker of *calculate* intervenes and blocks the creation of IG (or even C-to-C government, for that matter).

Word-initial three-member clusters in English are analyzed here as two overlapping IG domains as illustrated in (20). Since there are just two kinds of IG sequences, "ordinary" TR-clusters and exceptional, reanalyzed *sC* clusters, it follows that the only way in which the two can combine is *sTR*, i.e., where the first consonant (α) is /s/.[33] Since laryngeal features are shared under IG, the second consonant (β) must be voiceless and, in the case of plosives, unaspirated. That unaspirated plosive, however, has no independent laryngeal feature to share with the third consonant (γ, a sonorant), and as a result the usual effect of IG on voicing cannot manifest itself in the second pair. This derives the absence of sonorant devoicing in *sCC*.

[31] Even in Greek, a supposedly left-permissive language, there seems to be a systematic difference between the attested "exotic" initial sequences and the "accidental gaps", e.g., while /pt/ and /kt/ (both are traditionally classified as coda-onset) are well-formed word-initially, /tp/ and /tk/ ("bogus") are not. In fact, this is in sharp contrast to Arabic, where /tk/ appears to be one of the most frequent (syncope-created) word-initial occlusive-occlusive clusters, at least in the San'ani dialect (spoken in Sanaa, capital city of Yemen) according to Watson (2002: 76 fn. 18).
[32] See also Iverson and Salmons (1995).
[33] As it stands, the representation in (20) is unable to rule out nonexistent *sfR*.

(20)



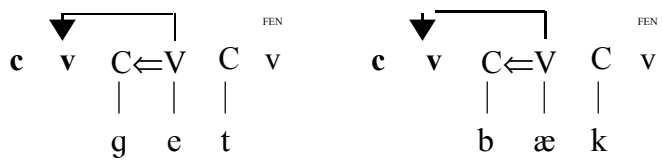
This concludes the discussion of consonant clusters, on the representation of which the last two sections spent some time and space. Let me emphasize again that they are not meant to provide comprehensive analyses. On the basis of t-lenition, they only suggest possible starting points for further research.

5.3.6 Word-final position

Since word-final consonant clusters were extensively discussed in the preceding section, and the behaviour of "unprotected" postconsonantal /t/ parallels that of single final /t/, only the latter is used here to illustrate the solution to the English cross-word puzzle.

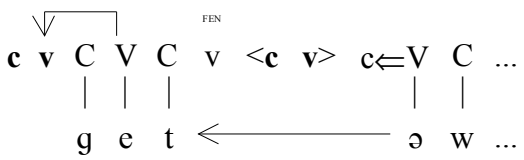
Recall (from Section 3.1.6 and elsewhere) that the stress-sensitivity of word-internal flapping vanishes when word-final t's (glottalised when utterance-final or followed by a consonant-initial word) are affected by a following vowel-initial word, irrespective of whether the vowel is stressed or not. However, it is only final t's that change according to the context, word-initial consonants are always strong. This latter case should be straightforward by now: the boundary-marker cannot be extraprosodic at the beginning of consonant-initial words by "Scheer's parameter". Consequently, consonants like the /b/ at the beginning of *back* and the /g/ at the beginning of *get* in (21) are licensed but not governed, i.e., are in a strong phonological position. The word-final /t/ (and, incidentally, the /k/, too) in (21), on the other hand, is unlicensed ungoverned, so it suffers consonantal lenition.

(21)*get back*



The spell-out rules concerning the allophony of /t/ can otherwise be shown to be U-level rules (cf. Nespor and Vogel 1986 and Chapter 4). It is argued in Section 5.3.1 above that V-to-C government proceeds on the melodic tier, so as long as the boundary-marker is extraprosodic at the beginning of vowel-initial words, the vowel's government is able to hit the preceding consonant. The relevant portion of the phrase *get away* is represented in (22), where the /t/ is followed by an unstressed vowel.

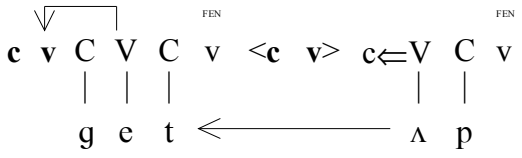
(22) *get away*



In (22) the schwa licenses its own empty c position on the CV-skeleton, rather in vain this time (cf. the discussion of hiatus-filling in Chapter 7). It can also govern, and since the boundary-marker is invisible to it along with the properly governable v position, it attempts to exert its governing potential upon a consonantal melody. The nearest target it finds is the /t/ at the end of *get*, which is therefore governed and undergoes vocalic lenition, i.e., tapping.

In *get up* (23), *get* is followed by a stressed vowel, but this time stress will not influence the state of affairs. The boundary-marker is underparsed in the same way as in (22), hence the government of /ʌ/ reaches the /t/ in the same way.

(23) *get up*



Crucially, the /t/ in (23) resides in a peripheral unit, i.e., a unit containing a morphosyntactic edge-marker, viz. the FEN. That explains why the stressed vowel is allowed to govern this /t/, as opposed to the one in, e.g., *atómic* (12b): the revised Antipenetration Constraint (4') only forbids government into non-peripheral units. The result is totally identical representations in (22) and (23).

5.4 A detour: l-allophony

The present thesis continually stresses the observation that vowel-initial words (and, supposedly, suffixes) exhibit special behaviour. This is not a novel idea. Itô (1989), for instance, based on the absence of "resyllabification" in VC#CV (≠> V.C#CV even if CC is an otherwise well-formed onset cluster)[34], concludes that "the universal aspect of syllable parsing is not onset maximization but onset satisfaction" (ibid: 222), and proposes the Onset Principle (Avoid $_{\sigma}$[v) (ibid: 223). In the present framework, this falls out naturally from the existence of empty c positions and the inherent ability of vowels to govern (for more detail, see Chapter 7): as shown in (22) and (23) above, vowels aim at exerting their governing potential even if this means intruding into a preceding word. Consonants, on the other hand, do not possess the inherent government charge; they only enter into government relations (C-to-C or IG) to satisfy the ECP of an enclosed empty v position.

[34] Of course, she means surface resyllabification as reported for Romance languages and discussed in Section 4.3. Still, the argument holds: no IG is contracted across a word boundary. She also mentions that VCC#V is never restructured as V.CC#V, which would not be too wide-spread anyway, considering the small number of overlaps between the sets of possible word-final and word-initial clusters.

To illustrate the power of melodically filled nuclei, let us take a brief overview of some of the dialectal variation that exists in English l-allophony. In RP, all l's followed by a consonant (other than /j/) or a pause are velarized (dark – cf. Section 4.4.2) whereas all l's followed by a vowel are clear irrespective of where the vowel comes from: whether it is part of the same word or in a following word (see, e.g., Jones 1966 or Rubach 1996). In contrast, all occurrences of /l/ are comparatively velarized in most dialects of American English (AmE), although some dialects except the foot-initial position and/or l's between high front vowels, e.g., *freely* (see, e.g., Ladefoged 1993 and Jensen 2000). Then, recall the AmE dialect dealt with in Nespor and Vogel (1986), referred to in Section 4.3. In this dialect, syllable-initial l's are clear (e.g., in *lap, Yolanda, vigilant* – cf. Nespor and Vogel 1986: 65), whereas all word-final l's are dark no matter what the following segment is (*pal, vigil, call, call Andy*). In the same dialect, l's remain dark when followed by a Class II suffix as in *vigilish.* A slightly different version is described in Halle and Mohanan (1985: 65): here there is a difference between the word-internal position and the compound boundary (*dealing, wheeling; a whale edition, the seal office* – all with clear /l/) and the phrase boundary (dark /l/ in *the whale and the shark, the seal offered a doughnut*).

Most probably, this dialectal variation stems from the fact that the rules of l-allophony choose different domains of application. In RP, it is the utterance, in Nespor and Vogel's case it is the word. In Halle and Mohanan's version, compounds may form a single p-word, and for some speakers an ungoverned licensed (= "foot-initial") /l/ may receive a different phonetic interpretation from that of a governed (= "foot-internal / ambisyllabic") one, resulting in the same stress-sensitivity as in t-allophony above. It is nevertheless remarkable that we only find variation before a following vowel, and rule domains cannot affect preconsonantal l's. If there is a distinction maintained between clear and dark /l/, then preconsonantal (and prepausal) l's are always dark. Also, the pretonic position quite reliably opts for the clear allophone: if there is a distinction maintained between clear and dark /l/, then (tautomorphemic) pretonic l's are always dark.

This observation receives additional support from a survey in Hayes (2000), in which ten native speakers of American English were presented words containing /l/ in various phonological positions, pronounced once with clear /l/ and once with dark /l/. The consultants were asked to judge the acceptability of both pronunciations for each word. The results echo the above conclusions: pretonic and word-initial /l/ sounds acceptable only if clear; preconsonantal and prepausal /l/ sounds acceptable only if dark. A kind of free variation is only attested in the foot-internal intervocalic position (heavily influenced by morphology, cf. *vigilish* above). Therefore, the preference for clear /l/ is strongest in words like *light, Louanne*, somewhat weaker in *gray-ling, gai-ly, free-ly; Mailer, Hayley, Greeley, Daley; mail-er, hail-y, gale-y, feel-y*[35], even weaker in *mail it* (most if not all of Hayes' informants seem to speak Nespor and Vogel's dialect, where the domain of l-darkening is the word), but weakest in *mail, bell,* and *help*. Notice that the native speakers' judgements show significant variation only in the classical "intervocalic ambisyllabic" environments.

This is readily explained by CV phonology's representations and the mechanisms outlined above: the government emanating from a vowel may be moderated by the boundary-marker, but a preconsonantal /l/ always occupies an unlicensed, weak position because of the following empty v position on the strict CV-skeleton, and because of the melody of the following consonant, which blocks any influence arriving from the right. Even if there is a following vowel potentially affecting the /l/, its government may be consumed by a boundary-marker (when the /l/ is initial) or simply wasted (when the /l/ is non-initial but the vowel is stressed). In addition, the theory predicts that postconsonantal /l/ (as in *butler*) is also dominantly clear: the following vowel's government is distracted from the /l/ by the empty v between the /l/ and the preceding consonant.

[35] The hyphens indicate morpheme boundaries.

5.5 Conclusion

This chapter has presented an alternative analysis of English plosive allophones, using t-lenition phenomena in GA as illustration. It exploits the suggestion made in Section 4.4.2 that V-to-C government proceeds on the melodic tier, and that it is the decisive factor to make the consonant lenite vocalically – whether the consonant is simultaneously licensed or not does not influence its phonetic interpretation. Thus flapping across word boundaries is also accounted for. The above discussion has addressed all the key issues identified in Section 3.2. The stress-sensitivity of word-internal lenition is due to the Revised Antipenetration Constraint, which makes reference to the distinction between peripheral and nonperipheral CV-units in such a way that it becomes irrelevant across words. Both word-initial and pre-stress positions are licensed but ungoverned, owing to the boundary-marker and the Antipenetration Constraint, respectively, which makes them phonologically strong. Foot-internal intervocalic consonants and codas, however, are weak, but for different reasons and that is why they (tend to) host different types of lenition. As elaborated on in Chapter 4, the visibility of the boundary-marker is governed by language- and process-specific factors (in English, it is always to be considered at the beginning of consonant-initial words according to the setting of "Scheer's parameter", but otherwise the domain of t-allophony is the utterance), ultimately resulting in function words behaving differently from lexical content words. The sections above also dealt with the behaviour of /t/ in consonant clusters and before syllabic consonants.

The analysis is fully consistent with a model of grammar in which a one-stage phonological component applies to the lexical representations of morphemes at the post-syntactic level. It makes no recourse to resyllabification or extrinsic rule ordering, in conformity with the basic tenets of GP, but, in contrast to previous GP analyses (Harris and Kaye 1990, Harris 1994), no cyclic derivation is needed, either. Government and licensing relations are simply established in response to, and abiding by, universal principles (e.g. the ECP or the inherent governing ability of nonempty vowels) and

language-specific parameter settings (e.g. "Scheer's parameter" or the Antipenetration Constraint[36]). Allophonic "rules" are in fact language-specific spell-out statements translating the emergent constellation of government and licensing relations (e.g., "ungoverned unlicensed") into melodic effects ("loss of the place element"). The cross-linguistic investigation of these statements outlines general tendencies, e.g. government effects vocalic lenition of consonants or silences empty v positions, whereas the absence of licensing leads to debuccalization or devoicing. On the basis of such recurrent patterns, it is possible to define government and licensing, as Dienes and Szigetvári (1999) did: licensing supports, but government spoils, the inherent properties of skeletal positions (cf. Section 2.1.3). Nevertheless, the definitions are undermined by phenomena like London glottalling (e.g. Harris and Kaye 1990), where both governed-licensed and unlicensed t̲'s are replaced by the glottal stop. In addition, there are a number of languages in which nothing happens to governed or unlicensed t̲'s. Such language-particular decisions support the view advocated here.[37]

The present account also observes the Projection Principle as interpreted in Section 1.2: governing and licensing relations are only added to the representation, never deleted or otherwise modified. This monotonous derivation does not involve the debatable "reduction" or "superimposition" operation familiar from Standard GP and from Section 1.2.

As a side-effect, the above analysis of t̲-allophony provides support for partitioning the skeleton into CV units. As it has been shown, VC phonology makes predictions which are not matched by the data. The innovations of CM+, however, have proven insightful, and are integrated into both the description of t̲-lenition and the tentative representations of consonant clusters.

One possible objection to the analysis is its failure to address the issue of the melodic representation of segments although it makes constant reference to interaction between structures on the melodic tier (V-to-C government, IG). The simple reason for

[36] It is proposed above that the Antipenetration Constraint be conceived of as a parameter deriving the difference between stress-sensitive and stress-neutral lenition systems.

[37] "[...] nobody can predict *when* a phonological event starts to be active [...] it is only when a phonological process indeed occurs that the theory makes predictions *where* and *how* it can and *where* and *how* it cannot operate, as well as *why* this is the case", as Scheer (2004: 247) points out.

this is the following. Element Theory (introduced in Section 1.6), as first presented in KLV (1985), has received fierce attacks from a number of otherwise faithful proponents of GP (e.g., Backley 1993, Scheer 1999a, Rennison 1998, 1999, or the so-called "Revised" Element Theory, developed at SOAS, London, in the 1990s), which has resulted in considerable dialectal variation even within the GP paradigm. There is no general agreement even concerning such elementary issues as universal vs. language-particular interpretation of phonological primes. The choice from the various versions is not at all trivial, and requires a substantial and thorough study of cross-linguistic melodic effects, which is beyond the scope of the present thesis. Nevertheless, I subscribe to privativity and full interpretability, the basic tenets, which ultimately imply an essential difference between natural classes in melodic complexity. The lack of a fully developed theory of melodic representations, however, does not necessarily invalidate or even undermine the claims of this chapter. Even if the "one-mouth" principle is correct and (more or less) the same set of primes takes part in the internal organization of consonants and vowels, V positions can be conceived of as inherently endowed with the ability to govern with their melodies irrespective of their complexity. As to C-to-C interaction, the exact nature of IG and C-to-C government is left unclarified since they also largely depend on melodic considerations, especially if one hopes to calculate complexity to determine the direction of government. Still, the claim that governing relations are established on the basis of the elemental make-up of the participants stays firm.

Therefore, as the present thesis is couched within GP, its model of consonant lenition subscribes to the basic tenets of Element Theory. The underlying form of /t/ is taken to be a relatively complex coronal plosive (whatever its exact elemental content is), whose gradual decomposition results in phonological expressions interpreted as glottalized or as a tap (in the spirit of the model making use of lenition trajectories as in, e.g., Harris 1994: 124). The relative complexity of the coronal stop is significant from the point of view of government: according to the Complexity Condition (Harris 1990, 1994; Section 1.6 above), governees cannot be more complex than their governors. This

is crucial in Harris and Kaye's (1990) and Harris's (1994) account (see Section 3.2.4) as well as the present analysis.

The other side of the coin concerns the issue of emptiness. Once communication between melodies is assumed, empty positions are announced totally empty (even void of, e.g., a "cold vowel"), and all phonetically interpreted positions must be melodically represented. That is, there are no phonetic objects which are mere realizations of phonologically empty positions but are opaque and have the same governing potentials as undoubtedly nonempty ones. One consequence is that (most) cases of vowel-zero alternation involve either underlyingly floating materials (as in Scheer 2004: 87ff and 552ff, for instance, who independently arrives at the same conclusion) or the insertion of "ambient" elements (as in Standard GP). In English, this specifically refers to the status of schwa. In the present context, it is considered to be an expression with an empty head and an **A** operator (KLV 1985, Scheer 2004 and p.c.) rather than the phonetic interpretation of an unsilenced empty nucleus lacking any melodic material (cf., e.g., Szigetvári 1999b)[38]. That schwa contains the element **A** receives independent support from its connection to /r/ (cf. Broadbent 1991 and Csides 1998).

In sum, the present analysis is capable of tackling the relevant data of t̲-allophony in a uniform fashion. Chapter 6 turns to the aspect of this phenomenon that is not treated above: lenition in two successive potential sites.

[38] Notice that Harris (1994: 181-182) is also forced to distinguish between genuine "empty" nuclei (represented as unheaded expressions consisting of the neutral element) and "stable schwas" (with the neutral element as a head), to be able to tell words like *dine* and *Dinah* apart. Considering the fact that the neutral element is usually thought of as a simple place-holder occupying all otherwise empty intersections of tiers and skeletal positions, drawing such a distinction smells like mere theoretical manoeuvring to avoid admitting that "stable schwas" are essentially different, that is, they are not void of melodic content.

Chapter 6: Weak and semi-weak phonological positions

This is the first of two shorter chapters discussing miscellaneous issues related to the main topic of the thesis. Although Chapter 5 has attempted to provide a comprehensive account of t-allophony in GA, there is one aspect of the phenomenon that we carefully but deliberately avoided: the so-called "Withgott-effect", i.e., the systematic absence of lenition in the third position of nonfinal dactyls, e.g. *mìlitarístic, Nàvratilóva*. This is the phenomenon to which we now turn our attention.

It has been observed by many (e.g., Withgott 1982, Jensen 1987, Harris and Kaye 1990, Burzio 1994, Steriade 2000, Jensen 2000, Davis 2003, 2005, just to mention a few) that in English (word-internal) dactylic sequences the unstressed position immediately following the foothead (that is, the first position after the stress) is more prone to reduce than the next syllable (that is, the second position). The Withgott-effect, or more precisely, the absence of lenition in American English in morphologically derived words like *mìlitarístic* (allegedly caused by the absence of lenition in the base *mílitàry*) has also been shown to be operative in that second position (Steriade 2000).

Vowel reduction appears to exhibit the same pattern. It is Burzio (1994: 113, footnote 14) who first pointed out that in English, foot-medial open syllables are affected by reduction to a greater extent than foot-final syllables[1]. In schwa syncope, too, a stronger tendency is attested for a first schwa to syncopate (e.g. *functionary* in the British English pronunciation) than for a second one (e.g. *functionary*).

In this chapter I propose to relate all these (apparently disparate) phenomena to a general distinction between weak and semiweak positions, originally introduced for Dutch in e.g. van Oostendorp (2000): the first unstressed position is weak, the second is semiweak. I argue, against Davis (2003), that there is no asymmetry between the behaviour of word-internal and final dactyls: the same strong-weak-semiweak pattern is detectable in both *militaristic/Navratilova* and *competitive/vanity*. In addition, pretonic unstressed syllables do not exhibit the same degree of phonological strength/weakness:

[1] Note that Burzio permits ternary feet, as opposed to a few others to be quoted below.

word-initially they are generally stronger (as in *potato*, with almost as much aspiration on the /p/ as on the first /t/) than medially (as in *Winnepesaukee*). Therefore, foot-based adjunction analyses, propagated in Withgott (1982), Jensen (1987, 2000), Davis (2003), etc. are inadequate either because they predict the same amount of aspiration in *Winnepesaukee* as in *potato*, or because they allow for a reduced vowel in a monosyllabic foot. Moreover, given the symmetrical behaviour of *militaristic* and *vanity* (claimed above), it is desirable that the two receive the same treatment, but the mechanism adjoining the third syllable of the dactyl to the right is clearly unavailable in *vanity*.

In my analysis I claim that, based on Ségéral and Scheer's (1999a) and Dienes and Szigetvári's (1999) definitions of government and licensing as two antagonistic forces, the theoretical framework used in the previous chapters is capable of expressing the relative weakness of the weak position of the dactyl, rather than the relative strength of the semiweak position. This way, it accounts for all the observations enumerated above without making reference to foot structure. The chapter is structured as follows. Section 6.1 introduces the concept of weak and semiweak positons in Dutch. The following sections argue for the relevance of the same distinction in English, with examples from t-lenition (Section 6.2.1, 6.2.2) including the Withgott-effect (Section 6.2.3) and from vowel reduction and schwa-syncope (Section 6.3). Then Section 6.4 provides the CV analysis, and Section 6.5 concludes the discussion.

6.1 Weak vs. semi-weak positions

The distinction between weak and semi-weak phonological positions was introduced for Dutch by van Oostendorp (2000: 147-8) to describe the propensity full vowels exhibit to alternate with schwa in stressless positions (basically, in free variation, the difference between reduced and unreduced pronunciations being one in style registers). What the Dutch data show is there are two types of unstressed position: one which is more prone

to reduce ("weak") and another with less frequent reduction ("semi-weak"). This is illustrated by the possible pronunciations of the Dutch word for 'phonology'.

(1) *fonologie* 'phonology'

very formal: [ˌfonoloˈɣi]

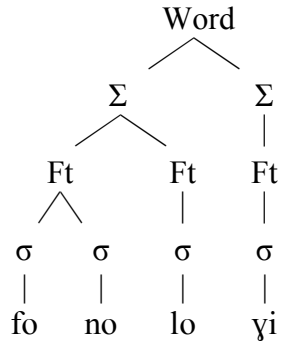
less formal: [ˌfonəloˈɣi]

even less formal: [ˌfonələˈɣi]

but: * [ˌfonoləˈɣi]

The conclusion to be drawn is that the syllable -*lo*- is more resistant to reduction: it is in semi-weak position. The usual analysis of this absence of reduction in the final syllables of word-internal dactyls makes reference to foot structure, as shown in (2), which is a reproduction of Figure (17) in van Oostendorp (2000: 148). Since feet are maximally binary, -*lo*- remains unfooted unless it erects a foot itself. As a result, both *fo*- and -*lo*- are footheads, and are therefore expected to be strong, as opposed to -*no*-, which forms the recessive position of the first foot. The two metrically prominent syllables (*fo*- and -*gie*) are different from -*lo*-, however, since they are heads of superfeet (Σ=superfoot).

(2)



In van Oostendorp's OT account, two separate constraints are proposed relating to footheads: one ensures that no reduction takes place in heads of feet (*fo*- and -*lo*-), the other bans reduction in heads of branching feet (and applies to *fo*- only). While

[ˌfonoloˈɣi] and [ˌfonəloˈɣi] do not violate either, and [ˌfonələˈɣi] violates the first one only, the unattested * [ˌfonoləˈɣi] would both violate the first constraint and contain an unreduced vowel in a non-foothead position.

6.2 Weak and semi-weak positions in lenition in English

6.2.1 Harris and Kaye (1990: 261)

As mentioned in Section 3.1.3, in their survey and GP analysis of t̲-lenition in New York City (NYC) English (tapping) and London English (glottalisation), Harris and Kaye (1990: 261) note the remarkable behaviour of words with two successive potential lenition sites, e.g. *competitive*. Here, two t̲'s are followed by an unstressed vowel, therefore both are expected to lenite. However, as Harris and Kaye observe, the second /t/ can only undergo weakening if the first one does so, too. This is illustrated for London glottalling in (3), but corresponding results are reported for tapping in NYC.

(3) *competitive*:
compe[t]i[t]ive
compe[ʔ]i[t]ive
compe[ʔ]i[ʔ]ive
*compe[t]i[ʔ]ive

Notice the parallelism between (1) and (3). Harris and Kaye are at a loss how to interpret this "'chain' of reduction". For them, lenition affects consonants trapped between two vowels which contract a licensing relationship. In this respect, there is no difference between the two t̲'s, as illustrated in (4): N_1, the stressed vowel, licenses both

N_2 (sandwiching α) and N_3 (sandwiching β). In fact, they use this as the motivation for lenition in both *political* and *sanity*.[2]

(4)

```
————————————————>
  ————————>
N1    O    N2    O    N3
 |    |    |    |    |
 x    x    x    x    x
      |         |
      α         β
```

However, the data can be reinterpreted as another manifestation of the weak vs. semi-weak distinction: there is a stronger tendency to lenite in the weak position (*compét<u>t</u>itive*), whereas the semi-weak position (*compétit<u>t</u>ive*) is more resistant to reduction.[3]

6.2.2 Difference between post-tonic and later positions

There is ample evidence that even in cases when only one lenition target is available, the behaviour of the immediate post-tonic consonant and that of the following position are asymmetrical. Several speakers of American English have reported that a /t/ immediately following the stressed vowel (e.g. *Italy*) *must* be a flap, a later /t/ (e.g. *sanity*) *may* be a flap. Hooper (1976), for example, claims that only post-tonic consonants are ambisyllabic, which is reflected by the fact that only such <u>t</u>'s are flapped (as in *kitty*) as opposed to intervocalic consonants not preceded by the stressed vowel (as in *serenity*, which contains an aspirated /t/ for Hooper). Others find that in words

[2] Harris (1994: 183) simply repeats the same representation, adding the labels "foot" and "super-foot" to the domains formed by the nuclei. In this way, he (ibid: 196-197) is again only able to account for the lenition of both targets in *competitor* (the first /t/ is foot-internal, the second is superfoot-internal), but not the lack of lenition in the second position.

[3] Notice that this statement goes straight against Szigetvári's (1999b: 47), which asserts that "in a word like *competitive* the two <u>t</u>'s are subject to the same type and degree of lenition".

like *capácity* or *éditor* aspiration is more acceptable than in *átom* or *glítter* (e.g. Selkirk 1982, Kreidler 1989: 110-111, Kenstowicz 1994: 69, Vaux 2002 and references therein). To quote Kahn (1976: 165 fn.17): "In some words which appear to be entirely on a par structurally with words like capital, failure to tap is not quite serious an affront to the American ear as the absence of flap usually is. Compare better, capital with marital. Even in the case of the latter word, however, /D/ is preferred greatly", whereas "[in immediate post-tonic position] as in better, unflapped /t/ is unnatural even in very careful speech" (ibid: 94). In a recent phonetic study, Patterson and Connine (2001), when listing the six possible phonetic environments of a medial /t/, only consider the immediate post-tonic position as the "flap environment", as in *water*, *party*, and list the intervocalic unstressed position (as in *parity*) under a separate heading, excluding it from the focus of their study. Vaux (2002) observes the difference between *consérvative* (with flapping) and *sédative* (without) – whatever causes this variation, it clearly affects the second unstressed position after the tonic.

For these speakers (and authors), this is a difference between weak and semi-weak positions: the later /t/ is in semi-weak position, and as such is more resistant to reduction. This manifests itself in free variation, but no such variation is found in the weak position.

6.2.3 The "Withgott effect"

Withgott (1982) was the first to highlight and analyse tap suppression in certain positions. She recorded that the /t/ is flapped in *capitalístic*, as expected, but aspirated in *militarístic, sanitisátion, monotonícity.* She pointed out that while *capitalistic* is morphologically related to *capital*, where the /t/ is already flapped, the untapped t's are all found in a derivative where there must be an untapped /t/ in the base due to stress on the syllable whose onset the /t/ is (*mílitàry, sánitìze, mónotòne*). She also argued that a cyclic analysis is not appropriate since aspiration (instead of lenition) is attested in words like *Meditérránean, Winnipesáukee, Navratilóva, abracadábra*, which are

morphologically independent. She proposed an adjunction analysis: stray syllables always attach to *adjacent* feet. Accordingly, *po-* in *potato* adjoins to the right while in words like *abracadabra*, after the localization of footheads, there remain two stray syllables inbetween; the first is adjoined to the left, while the second adjoins to the right. In this way, the second unstressed syllable following the stress becomes foot-initial, therefore strong. This is a modification of Hayes (1982), where both stray syllables are assumed to adjoin to the left.

Jensen (1987) pursues the same idea as Withgott[4]. He brings supporting arguments from native intuition (informants were asked to divide words like *abracadabra* into two parts), and verifies the results with instrumental measurements of the duration of stop release in the same words, only to underpin Withgott's intuitions. He concludes that the third syllables of such words are footheads, since only foot-initial voiceless plosives are aspirated. The stray syllable adjunction rules he assumes are explicitly formulated in Jensen (2000: 210), where he derives the difference between *capitalistic* and *militaristic* in terms of a cyclic derivation of stress and foot structure. In the first cycle in *capitalistic*, the only possibility is for (*capital*) to be assigned a dactylic foot (enclosed in the parentheses), which is preserved in the second cycle, yielding (*capita*)(*listic*). In both instances the /t/ is foot-internal. In *military*, however, two feet are produced in the first cycle ((*mili*)(*tary*)). In the second cycle, the foot (*ristic*) is formed on the right and stray *-ta-* adjoins to the right, giving (*mili*)(*ta*(*ristic*)).

Although it has the same effect, Withgott's and Jensen's solution is just the opposite of van Oostendorp's in (2) above.

Steriade (2000: 322-326) also addresses the problem, though she approaches it from a completely different angle. She claims that paradigm uniformity (PU) is at work here. PU promotes the invariance of some sound property within a paradigm, and is defined as given in (5).

[4] He does not refer to Withgott, though.

(5) Paradigm Uniformity

All surface realizations of μ, where μ is the morpheme shared by the members of paradigm x, must have identical values for property P. (Steriade 2000: 313)

Tap suppression in words like *militaristic* is a PU effect, Steriade claims: it is the paradigmatic extension of the unflapped stop of *military*, more precisely, of the [extra-short closure] feature the flap does, but the stop does not, possess. To show that the Withgott-effect is systematic, Steriade presents the results of a survey she carried out with 12 speakers of American English, who were asked to read out the following (often nonce) words.

(6) a. Bases: *positive, primitive, relative, negative, voluntary*
Derivatives: *positivistic, primitivistic, relativistic, negativistic, voluntaristic*

b. Bases: *rotary, fatal, fetish, totem, notary*
Derivatives: *rotaristic, fatalistic, fetishistic, totemistic, notaristic*

Derivatives in *-istic* are expected to display stem invariance effects since the morphological operation producing them is highly productive, and they are (fully) compositional. In the bases in (6a), speakers differ as to whether *-ive* and *-ary* are stressed – consequently, whether the /t/ can be flapped. The quality of the /t/ in the bases is predicted to determine the one in the derivatives. In the words in (6b), however, all intervocalic t's should be tapped, being followed by stressless vowels. In the survey the words in (6a) were mingled with the words in (6b) to minimize the influence of similar words on each other, and also to check whether the informants were not producing artificially untapped pronunciations. The results showed virtually no exceptions to base-derivative correspondence, i.e., whenever there was a tap in the base in the examples in (6a), there was a tap in the derivative, and whenever there was an untapped /t/ in the base, it was unchanged in the derivative.

In monomorphemic strings PU is irrelevant, and, as Steriade observes, t̲'s are generally tapped in unstressed position in words like *meri̲tocratic, hema̲togenesis, peri̲tonitis, hema̲tocystic*. She insists that *Mediterranean* is a unique underived form in which the tap is suppressed; her explanation is that the orthographic geminate <rr> is interpreted by speakers as an indication of secondary stress on the preceding vowel. Unfortunately Steriade does not comment on Withgott's other examples, which do not contain orthographic geminates in the relevant position at all (cf. *Winnipesaukee, Navratilova, abracadabra*). Vaux (2002) adds *Vinatieri* to the list, and cites *lollapalooza* from Davis (2001). Among others, Jensen (2000) and Davis (2003, 2005) insist that this reflects a regular pattern, which goes against Steriade's analysis.

What is of really high relevance to the present discussion, however, is what Steriade remarks in endnote 4: tap suppression does not obtain in syllables that directly follow the tonic, as in word pairs like *sta̲tístic – sta̲tistícian*; in the second item in these pairs a tap usually appears and, generally, there are very few instances of non-tapped t̲'s in the V_v context: "[...] constraints that induce tapping are more stringent (i.e. more highly ranked) in the immediate post-stress position than elsewhere. PU effects surface only when the tapping constraint is weaker." In light of the foregoing discussion, this remark can be interpreted to argue that examples of tap suppression (whether or not they are manifestations of PU effects) are only found in the semi-weak position, irrespective of morphological structure.

Section 3.1.3. mentions the claim made in Davis (2003) that there is an asymmetry between final and nonfinal dactylic sequences. Based on data from flapping, aspiration and expletive infixation, Davis finds that only nonfinal dactyls contain a strong third syllable. He ignores the variation described above and states that in words like *cápital, serendípity, chárity* the /t/ is "clearly flapped" (ibid: 278). He also cites Van Dam and Weaver (2001), who show that the voiceless stops /p/ and /k/ at the beginning of the final syllable of words like *América, Connécticut, Oédipus* are lightly aspirated and have neither the degree of aspiration that accompanies a foot-initial voiceless stop nor the aspiration that is attested in nonfinal dactyls like *Medi̲terranean*. He concludes

that this is due to the fact that the stops in words of the *capital* and *America* type are foot-internal.

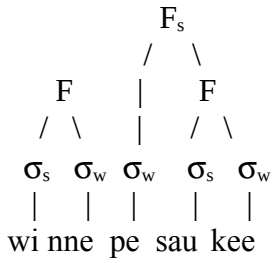
In contrast, in nonfinal dactylic sequences the third syllable is adjoined to the right, and therefore it is not foot-internal but foot-initial (in the same way as in Withgott 1982 and Jensen 2000). He argues that both Jensen (2000) and Pater (2000) independently contend that a voiceless stop in that position is indeed aspirated in American English, and again refers to Van Dam and Weaver's (2001) study considering the voice onset time of the voiceless stop at the beginning of the third syllable in the words *Wìnnepegósis*, *Mèditerránean*, and *Nèbuchadnézzar*. They found that these stops had an average voice onset time of more than 50 milliseconds, which is almost as aspirated as pretonic stops, and definitely more aspirated than voiceless stops at the beginning of stressless syllables immediately after the stressed syllable in words like *múppets* and *móccasins*, with very short voice onset times, less than 20 milliseconds on average for the non-coronal stops (the coronal flap is even briefer).

Expletive infixation, i.e., the infixation of an emphatic element like *fuckin', frickin, bloody, bloomin'*, etc. (described in, e.g., Aronoff 1976) can also be used to detect foot boundaries. The generalization regarding expletive infixation is that the expletive occurs before the foot boundary, e.g. *po-fuckin'-tato, Ne-fuckin'-braska*. McCarthy (1982) observed that words like *Wìnnepesáukee* show variation with respect to this process: both *Winne-frickin-pesaukee* and *Winnepe-frickin-saukee* are possible, which suggests that both *-pe-* and *-sau-* are footheads. McCarthy's observation regarding the variant forms appears correct and robust, as Davis (2003) says, adding *mili-fuckin'-taristic* and *milita-fuckin'-ristic*. However, applied to *càpitalístic*, only the form *capita-frickin-listic* is judged to be acceptable, with the expletive after the foot-final third syllable. The form **capi-frickin-talistic* appears illicit, especially if the flapping of the /t/ of the original third syllable is maintained.

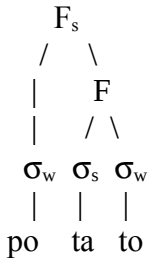
Notice that the data regarding expletive infixation is only relevant as an argument for the different footing of *militaristic* and *capitalistic*, which Davis (2005) sees as a PU effect: like Steriade, he claims that the flap in *capitalistic* can be accounted for by paradigm uniformity with *capital*. However, unlike Steriade, he argues that the

expected regular pattern is the one found in *mili<u>t</u>aristic*, that is, aspiration in the third syllable of a word-internal dactyl due to its foothead status, and it is *capi<u>t</u>alistic* that exhibits flapping because of uniformity of foot structure with *capital*. This explains aspiration in underived words like *Navratilova*, too, which was left unanalysed by Steriade. A weakness of Davis's (2005) treatment of foot structure and aspiration is illustrated in (7): in the representations taken from his paper, the superfoot is the same whether it is word-initial (as in *potáto*, see (7b)) or not (as in *Wìnnepesáukee*, see (7a)), although undoubtedly aspiration is stronger word-initially.

(7) a. *Wìnnepesáukee*



b. *potáto*

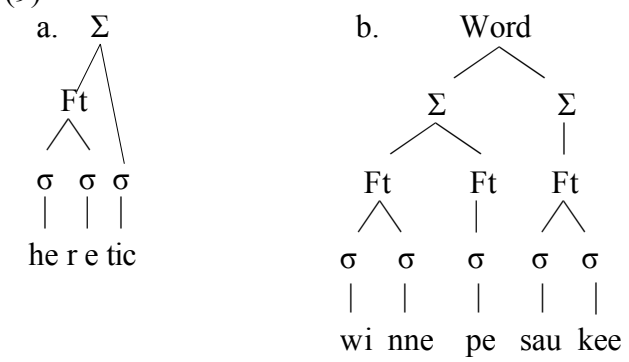


Of course, expletive infixation is unable to test the assumed asymmetrical behaviour of final and nonfinal dactyls, since infixation is not possible when no foot boundary follows, as in *vanity*. What remains as an argument for Davis is <u>t</u>-allophony, which is considerably weakened by the data cited in the preceding section (6.2.2), suggesting that there is not such an asymmetry. So much so that Anderson and Ewen (1987: 83) propose a similar superfoot-structure for words like *hére<u>t</u>ic, aríthme<u>t</u>ic*, where the underlined /t/ is aspirated rather than tapped. Their figure 2.97 is reproduced below in (8). It shows that within trisyllabic feet like *heretic*, they assume that since metrical structure in English is binary, the stressed vowel (in *he*-) first forms a so-called subfoot (*here*-, indexed with 1), and then a superfoot (indexed with 2). The consonant immediately following the stressed vowel (/r/) is foot-internal and therefore ambisyllabic (i.e., belongs to two syllables at the same time, that is why it is surrounded by brackets in (8)), whereas the next consonant (/t/) is exclusively syllable-initial because it is shielded from becoming ambisyllabic by the bracket closing the subfoot.

(8) $_2[\ _1[[he[r]e]]_1\ [tic]]_2$

The basic difference between Davis' *Winnepesaukee* in (7a) and Anderson and Ewen's *heretic* in (8) is in the direction of adjunction for the formation of the superfoot. In this respect, Anderson and Ewen's model resembles van Oostendorp's analysis of Dutch *fonologie* in (2) above; in (9), similar representations are given corresponding to *heretic* (9a) and *Winnepesaukee* (9b).

(9)



(9a) is the arboreal representation of (8), while (9b) is a total analogue to *fonologie* (2). In both, the third syllable of the dactyl is adjoined to the left; the difference is that in (9b) it forms a foot itself. This is due to the fact that Anderson and Ewen make use of the distinction between ambisyllabicity and absolute onsethood to derive the difference between the /r/ and the /t/, whereas to van Oostendorp this option is unavailable, all single intervocalic consonants being unambiguously parsed as onsets for him, he could only make *pe* stronger than *nne* by assigning it to a foot. This is the only way we can keep up the parallelism between final and nonfinal dactyls (viz., the third syllable is adjoined to the left) *and* make the /p/ in *Winnepesaukee* foot-initial (to account for its aspiration). The objection arises then, that if *pe* is a foothead, which it is in (9b), how is it able to reduce its vowel to a schwa? This is not normally expected from footheads. Notice that the final syllable of *heretic* should retain its non-foothead status for the same

reason: compare it to *hesitate*, for example, whose final syllable *must* erect a foot and consequently its vowel cannot reduce, and the initial /t/ of *-tate* is mandatorily aspirated and/or untapped. The vicious circle has closed: we are back with *-pe-* in *Winnepesaukee* adjoined to the left without projecting an intervening foot level – complete analogy with (9a), i.e., no difference between final and nonfinal dactyls, but then why is their third syllable stronger than the second?

We conclude that a foot-based analysis is inadequate to account for the asymmetry between weak and semiweak positions on the one hand, and for the symmetry between final and nonfinal dactyls on the other. If one gets rid of ambisyllabicity as a theoretical device, neither possible adjunction analyses are fully satisfactory: they either predict the same amount of aspiration in *Winnepesaukee* as in *potato* and/or *hesitate*, or they allow for a reduced vowel in a monosyllabic foot.

6.3 Weak and semi-weak positions in vowel reduction and schwa syncope

This section aims to provide an attempt at finding the English analogues of Dutch *fonologie*, that is, traces of the weak-semiweak distinction in vowel reduction (to schwa or zero). It is Burzio (1994: 113, footnote 14 – also cited in van Oostendorp 2000) who first pointed out that in English, foot-medial open syllables are affected by reduction to a greater extent than foot-final syllables[5]. That is, for a word like *Tatamagouchi*, the pronunciation *(tætəma)gouchi* is preferable to *(tætamə)gouchi*, i.e., if one of the two unstressed vowels of the first foot (parenthesized) remains unreduced, it is more acceptable for the first (underlined) to reduce. The words *(rigama)role, (panama)* are claimed to behave analogously. This observation is totally analogous to the Dutch example in (1) above, and is straightforwardly interpretable as the tendency in semi-weak position for vowels to be more resistant to reduction. Moreover, if *panama* is

[5] Recall that Burzioallows for ternary feet, in contrast to most authors referred to in the discussion above.

analogous to *Tatamagouchi*, then this is additional evidence of the absence of asymmetry between word-internal and final dactyls, argued for in the previous section.

Burzio (ibid.) apparently finds a parallel situation as regards vowel syncope, that is, schwa-deletion: in *memorization*, says Burzio, the first foot contains two schwas in a row, out of which only the first can undergo syncope, i.e., *(mem'ri)zation* is a possibility, while **(memor')zation* is not. Before arriving at too hasty conclusions, however, one must recognize that *memorization* is not the most fortunate example since, besides the well-known rarity of immediate pre-stress syncope in other than word-initial syllables, which renders the deletion of the schwa of *-ri-* (followed by primary-stressed *-za-*) highly improbable, the segmental context (*r_z*) does not support its deletion, either, since the consonant following the potential site is an obstruent.

To avoid such factors inhibiting vowel deletion, I carried out a survey with words like *functionary* and *nationally*, where a sequence of two unstressed (therefore syncopatable) vowels appears in the right segmental context (i.e., $C\underline{v}S_1\underline{v}S_2v$, where C is less sonorous than S_1, which is in turn less sonorous than S_2; S=sonorant consonant, and the third vowel is also unstressed). I used EPD[6], LPD, and native informants to find out about the preferences of schwa deletion in such words. Unfortunately, there are not much more than 60 words that qualify for the present purposes, and this small number of examples is made even smaller by the fact that the majority of the sample consists of derived words, in the case of which Paradigm Uniformity (PU) effects can influence the choice of pronunciation (see below). Also, the application of syncope is heavily influenced by word frequency (cf. Hooper 1978): less frequent words strongly resist it even if all the phonological conditions are met, and natives are unable to judge nonsense words. Still, there remain a few examples in which the weak-semiweak distinction is able to manifest itself in spite of the morphological pressure, e.g. *confectionery* and *functionary* (-ʃnəri being more frequent than -ʃənri).

In a number of clearly derived words, PU effects (see (5) and the discussion in Section 6.2.3 above) are detectable: there is a tendency for syllable peaks to be preserved in the derivatives, cf. (10). In the first column of the chart I use *nationally* to

[6] Thanks to Péter Szigetvári for making it available for online browsing.

illustrate the behaviour of adverbs formed from -*al* adjectives by affixing -*ly*. In the second column, I use *cautionary* to represent words ending in -*ary*.[7] The vertical line denotes the relevant morphological boundary.

(10)

national	ˈnæʃnəl	*caution*	ˈkɔːʃən
nationally	ˈnæʃnəl|i	*cautionary*	ˈkɔːʃən|ri
	*/? ˈnæʃənl|i		*/? ˈkɔːʃn|əri

As the chart shows, *nationally* prefers deleting the first schwa, most probably because it can also be deleted in the base *national*. All the -*ly* words that I was able to check follow this pattern. In the second column, however, *cautionary* demonstrates how the great majority of -*ary* words that I checked behave.

It is evident that schwa syncope in English is a complex phenomenon, influenced by a host of nonlinguistic or nonphonological factors. In addition, the quality, more specifically the sonority distance, of the consonants flanking the syncope site may affect a vowel's proneness to delete, as claimed in, e.g., Zwicky (1972). Nevertheless, a generalization can be made to the effect that, in words of less transparent morphological structure at least, the weak-semiweak distinction seems to be justified.

6.4 Analysis

It has been demonstrated above that previous foot-based accounts of the configurational aspect of lenition fail to properly describe the strong-weak-semiweak tripartite distinction.[8] In what follows I present an alternative analysis, using the Strict CV framework as introduced in the thesis. I heavily rely on Ségéral and Scheer's (1999a) and Dienes and Szigetvári's (1999) definitions of government and licensing as

[7] The option of syllabic consonant formation is disregarded.
[8] The nonconfigurational aspects are word frequency, PU, and the like.

two antagonistic forces. I claim that, crucially, it is not the third syllable of dactyls that is *stronger* than the second, it is the second that is *weaker* than the third. To explain this, I propose that once stress assignment designates a vowel to be (primary or secondary) stressed[9], at least the following three features ensue: it (i) falls under the rubric of the Revised Antipenetration Constraint ("Stressed vowels are unable to govern into non-peripheral units" – (4') in Section 5.2); (ii) resists the Proper Government (PG) emanating from a following filled vowel, and instead (iii) distracts the licensing charge of the following vowel[10]. This results in grave consequences for the immediately following CV-unit. On the one hand, the vowel residing there will never be able to properly govern another one, and therefore its PG will always hit its C; on the other hand, its licensing charge is diverted from the C by the stressed nucleus: the consonant is expected to exhibit a strong tendency to undergo vocalic lenition. Meanwhile, the obligation to license the metrical head exhausts it, so much that it becomes weaker, i.e., easier to reduce or delete. In contrast, later unstressed vowels are free to properly govern either their onsets or the preceding unstressed vowel (reducing it even further), and also to license their onsets. Therefore, such nuclei are stronger, and their onsets are both licensed and governed: these C's are able to lenite, since they are governed, but they do so in such a way that licensing, supporting their melodic expression, reaches them as well.

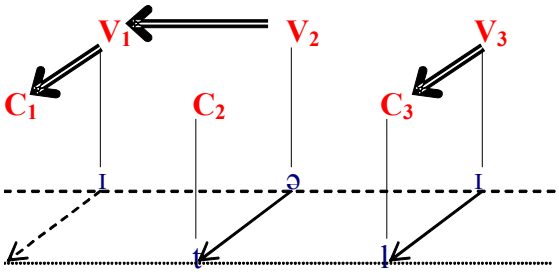
In sum, the weak-semiweak distinction observed for consonants is due to the difference between the governed unlicensed and governed licensed position: the former is weaker than the latter. In the case of nuclei, on the other hand, the distinction results from the fact that the unstressed vowel immediately following the stressed one is forced to deplete its licensing potential via a marked relationship, V-to-V licensing on the nuclear projection. The novelty of this analysis lies in its emphasis on the relative weakness of the weak position of the dactyl, rather than the relative strength of the semiweak position.

[9] A possible model of how that happens is sketched out in Szigetvári (1999b) and Scheer (2004: 613ff).
[10] V-to-V licensing has been proposed to be responsible for long vowels (Szigetvári 1999b) or for vowel length alternations and to facilitate the survival of schwa (in French at least) (Scheer 2004). The one sketched out here is yet another possibility, viz., that its target is dependent on stress relations.

In (11), (12) and (13) below, a possible representation of governing and licensing interactions is sketched out. As Chapter 5 proposes that V-to-C government takes place between melodies, and it has long been believed that V-to-V relations are contracted on the nuclear projection, but V positions only have one shot of government and licensing each, which results in the complementary distribution of V-to-V government (i.e., PG) and V-to-C government on the one hand, V-to-C licensing and V-to-V licensing on the other, it is apparent that this complex network of lateral relations existing on various levels of representation (i.e., tiers) can be best modelled in three dimensions. This is demonstrated by the representations of *Italy* /ˈɪɾəlɪ/ (11), *...petiti...* (from *competitive*) /...ˈpʰeɾətʰɪ.../ (12) and *Italy* /ˈɪtˀlɪ/ (13). The upper level is the CV-tier (typed in grey): this is where skeletal positions communicate. The CV-tier itself is however made up of two tiers, the C-tier (where C positions are adjacent, in, e.g., consonant harmony systems) and the V-tier (the former "nuclear projection", where V positions interact, irrespective of the enclosed C's). The melodic tier (typed in black) is structured analogously: it is composed of vocalic melodies (along the broken line) and consonantal ones (along the dotted line). Government applies on either the V-tier or the melodic tier, whereas licensing is exclusive to the CV-tier. The skeletal positions are in cases other than floaters linked to their respective melodies by vertical association lines, as usual in autosegmental representations. These association lines are of utmost importance since they ensure that a skeletal slot and its melody are in fact one and the same object, together constituting the segment. Thus, PG is associated with the V-tier by definition, but the same governing relation may hit a consonant on the melodic tier (at the lower level) and effect lenition[11]. Notice that all binary relations in such a model are strictly local, even PG.[12] [13]

[11] This model resembles Scheer's (2004: 243ff, to appear) in distinguishing between two levels of phonological interaction (HIGH and LOW in Scheer's version), although the conception of the exact "geographical" distribution of government and licensing relations differs in several respects.
[12] Unbounded stress or harmony domains will still involve relations which are not local.
[13] Another way of "regaining" locality is demonstrated in Szigetvári (1999b: 109): all lateral relations apply between two adjacent skeletal units.

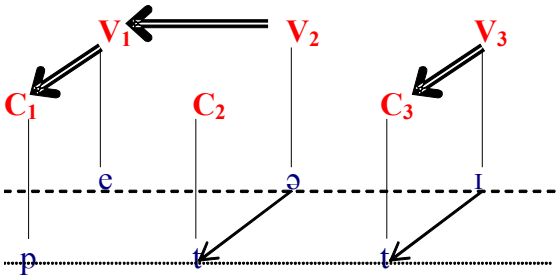
(11) *Italy* /ˈɪɾəlɪ/



In *Italy*, when the second vowel (V_2) is not properly governed and is therefore pronounced, the final vowel (V_3) governs the melody of the preceding consonant (/l/) (single arrow), but it also licenses it (double arrow): the consonant finds itself in a governed licensed position, that is, it is semiweak. C_2, however, is weak since it is governed only; V_2's licensing is consumed by V_1, the metrical head, which in turn licenses C_1, and also attempts to govern it, but as it is empty, this government can only manifest itself in connected speech when the preceding word ends in a consonant, e.g., *hate Italy* (with the underlined /t/ potentially tapped).

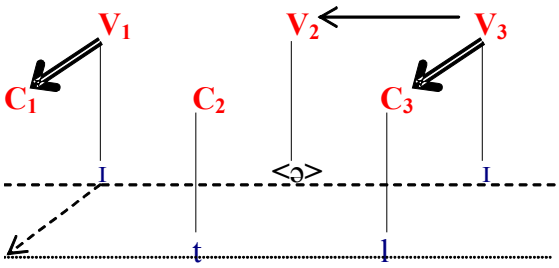
Now consider the example of the strong-weak-semiweak distinction in tapping, the relevant portion of *competitive* /...ˈpʰɛɾətʰɪ.../ in (12). The /p/ is licensed only (and aspirated as a consequence), since according to the (Revised) Antipenetration Constraint, stressed vowels cannot govern into non-peripheral units. Therefore, this is a strong phonological position. C_2, however, is weak, for the same reasons as the /t/ in *Italy* in (11), and C_3 is semiweak because it receives both government and licensing: it is expected to vacillate between a tap, an unaspirated voiceless plosive, and an aspirated one.

(12) *...petiti...* (from *competitive*) /ˈpʰɛɾətʰɪ/



Finally, *Italy* /ˈɪtˀlɪ/ illustrates PG (between V_3 and V_2), which results in the underparsing of V_2's melody (indicated by the empty box). Governed vocalic positions are phonetically uninterpreted and are deprived of all their governing and licensing capacities. Therefore, C_2 is ungoverned unlicensed, and as such is expected to lenite consonantally, i.e., undergo glottalisation.

(13) *Italy* /ˈɪtˀlɪ/



The advantage of this model over previous ones is that it does not only cover all the observations enumerated in the preceding sections, but it does so in such a way that it makes use of notions and principles (government, licensing, the Antipenetration Constraint, etc.) which have been independently motivated throughout this thesis and elsewhere, and it conforms to the fundamental tenets of GP and CV phonology (e.g. locality and the denial of ambisyllabicity).

6.5 Conclusion

This chapter makes two fundamental claims. On the one hand, it is extensively argued for that the distinction between weak and semi-weak phonological positions seems to be justified in English, too, besides other Germanic languages like Dutch, which exhibit the same type of stress-sensitive pattern of lenition. On the other hand, it is shown what weaknesses earlier foot-based analyses suffer from, and how the CV framework, more specifically, the notion of government and licensing as two relations of opposing effect, is capable of expressing the tripartite distinction between strong (licensed ungoverned), weak (unlicensed governed) and semiweak (licensed governed) phonological positions. The fourth logical possibility, that of an unlicensed ungoverned position is also weak but triggers lenition along a different trajectory.

This models also explains what happens in word-initial unstressed syllables, traditionally analysed as degenerate feet or as unstressed syllables adjoined to the right, e.g. in *potato*. Since such syllables display a hybrid-like behaviour, with a strong consonant but a weak vowel, neither of the two proposals describe them properly. In contrast, it falls out naturally from the present analysis that this is indeed the expected state of affairs: although the vowel is unstressed and therefore reduced, its licensing charge is not diverted from its onset consonant as there is no stressed vowel to the left; its government, however, avoids the C because the silencing of the boundary-marker is of higher importance. Consequently, the /p/ of *potato* finds itself in a licensed ungoverned, i.e. strong, position in the same way as the pretonic /t/.

Chapter 7: Empty onsets

7.1 What's wrong with vowel-initial syllables – introduction

The whole of Chapter 5, and Section 5.4 in particular, emphasizes the special role of vowel-initial words. In this chapter we extend our scope to all vowel-initial syllables, or, in Strict CV/VC terms, to empty C positions. Recall Itô's (1989) conclusion that "the universal aspect of syllable parsing is not onset maximization but onset satisfaction" (ibid: 222), and the Onset Principle cited in (1).

(1) The Onset Principle (Itô 1989: 223)

Avoid $_{\sigma}$[v

The question arises why onsets enjoy such a privilege. It is well-known that codas are universally highly marked, a number of languages do not even allow for syllable-final consonants, others impose serious restrictions on them. In CV phonology, codas are defined as consonants followed by an empty nucleus, and this theoretical framework accounts for coda effects with reference to the marked status of empty V positions ("V slots in the phonological skeleton aim at being pronounced" – Szigetvári 1999b: 62) and the weakness of unlicensed C's. Onsets, however, are frequently mandatory, or at least desirable, as expressed in (1), although this does not follow logically from the definition of consonantalness ("...if nothing intervenes a C position will stay silent" – Szigetvári ibid: 62). If C positions are inherently mute, what explains the universal markedness of this muteness?

The answer lies in the definition of onsets: they are C positions followed by an ungoverned vowel, i.e., a nucleus endowed with full governing and licensing potential. Recall from Section 5.4 the claim that vowels aim at exerting their governing capacity even if this means intruding into a preceding word. The ungrammaticality of empty C's

followed by an able V stems from the fact that an empty C position is not associated with any melody for the vowel to govern, that is, such vowels are frustrated by not finding a way to wield their influence. This serves as an additional piece of evidence that there is an interaction between melodies, which we refer to here as government.

7.2 Hiatus versus vowel-initial words

There are two possible positions for empty c's: word-medial (in which case the situation is a hiatus)[1] and word-initial (defining a vowel-initial word). This section explores what Strict CV phonology has to say about these two locations and about the relationship between them.

Hiatuses are dispreferred in languages, and various repair strategies exist to avoid them (including vowel elision, hiatus filling, coalescence). The cross-word investigation of elision under hiatus (e.g. Casali 1997) reveals that in a V_1V_2 sequence there is an overall preference for V_1 elision. Although the type of juncture flanked by the vowels can be a conditioning factor (e.g., when a vowel-final root or word receives a vowel-initial suffix, either V_1 or V_2 can be elided, the former in the default case, the latter if base identity overrides it, the choice being either language-specific or the same language may exhibit both types), crucially there is no language on record that systematically elides V_2 in *all* environments.

This universal asymmetry between the two members of hiatuses suggests that the force responsible for eliding one of them is a right-headed relation. Therefore, (V-to-V) government qualifies as a possible candidate: its direction is in most cases (if not universally) right-to-left, and its effect is the demolition of the target's inherent loudness[2]. Also, recall that in Strict CV phonology it is exactly the V_1cV_2 configuration

[1] Throughout the discussion, long vowels and diphthongs are ignored – they may be manifestations of V-to-V licensing, as claimed in Szigetvári (1999b), although this is in conflict with the analysis of weak positions in Chapter 6, where V-to-V licensing is proposed to hold between a metrical head and a following nucleus.

[2] This is not a novel idea: Charette (1991: 24), e.g., explicitly claims that adjacent nuclear positions enter into an interconstituent government relation.

where V_2 is frustrated by not being able to govern since the melody of the consonant is missing. Thus I propose that the two vowels of a hiatus enter into a right-to-left governing relationship. This proposal is in sharp contrast with previous CV/VC analyses (e.g. Szigetvári 1999b: 72-73), which claim the absence of V-to-V communication in hiatus. Notice, however, that were there not such a relation, no reason could be found for favouring the vowel to the right rather than the one to the left of an empty c.

It is claimed in Chapters 5 and 6 above that government may proceed on either the V-tier (the "nuclear projection"; e.g. Proper Government) or the melodic level (e.g. V-to-C government). It may be the case that V-to-V interaction in hiatus is similarly associated with either of the two, determined by a language-specific or other choice: if V_2 governs V_1 on the V-tier, V_1's ability to support melody is destroyed altogether, so hiatus is expected to be resolved by vowel deletion (or gliding, as in Szigetvári's (1999b: 82-83) analysis of *social, medial*, and French glide-formation in words like *lié*), since this is the usual effect PG has on its targets. If, on the other hand, it governs V_1 melodically, government attacks V_1's melody directly, and therefore it is forced to decompose (e.g. diphthongs undergo monophthongization). One way of avoiding either of these two unpleasant situations involves letting V_2 govern a consonant, i.e., inserting or spreading melody into the sandwiched c position: then V_2 governs the hiatus-filler melodically, which is therefore required to be as vowel-like as possible (a sonorant, ideally a semivowel).[3] [4]

Strict CV phonology predicts that a vowel-initial word behaves differently from a hiatus, as the beginning of the word is signalled by the boundary-marker. The left edge of the word contains a cvcV sequence, where the V has the job of governing the empty v of the boundary-marker in order to silence it. Consequently, the prediction is that empty c's are tolerated more at the beginning of the word than medially: as long as the boundary-marker is not extraprosodic, it attracts the PG of the initial filled vowel,

[3] This is similar to Szigetvári's (1999b: 105) analysis of French liaison: he claims that a floating consonantal melody is linked to an empty c position if it is governed. Since government spoils the inherent properties of its target, and C positions are inherently mute, government induces their phonetic interpretation.

[4] It was brought to my attention by Miklós Törkenczy that the behaviour of hiatus in Margi, described (for a completely different purpose) in Kenstowicz (1994: 321), falls out naturally from my analysis.

irrespective of other parameter settings concerning vowel sequences, and no filled V_1 is present to induce the melodic interpretation of the c position[5]. This prediction is born out by the data: languages with vowel-initial words but no internal hiatus are larger in number than languages with the opposite preferences (cf., e.g., the extensive survey of Smith 2002).

Another prediction concerns the so-called permissive languages (cf. Section 4.1), where the boundary-marker is invisible to phonotactics. From the above discussion it follows that in that case the government of initial filled vowels is not diverted from their empty onsets, which are thus more likely to lose their inherent muteness and be pronounced. That is, in permissive languages no or fewer vowel-initial words are expected. Whether this prediction is empirically supported still needs to be examined, but in Arabic, a rather uncontroversially permissive language, no word may start with a vowel (cf. e.g. Watson 2002: 65ff).

In sum, in the CV analysis vowel-initial words and medial hiatuses are treated differently, and as a consequence languages are expected to have the choice of tolerating one or the other more easily. This claim seems to be verified, as shown in (2), which is a tabular representation of language typology concerning empty onsets, i.e., syllable-initial vowels (abbreviated to [V).[6]

[5] Of course, in connected speech the boundary-marker may be extraprosodic within the limits of some prosodic constituent, and then hiatus resolution applies in the same way across words as internally, that is, hiatus resolution can be bounded within some domain.

[6] The chart draws heavily on Smith (2002) as well as on responses I received to a LinguistList query; special thanks to Joaquim Brandao de Carvalho, Rod Casali, Daniel L. Everett, Jennifer L. Smith, and Jan-Olof Svantesson.

(2)

[V

		word-initially: yes	word-initially: no
word-internally	yes	English, Hungarian, etc.	Arapaho, Guhang Ifugao, Hausa, Guaraní, Tabukang Sangir
	no	European Portuguese, Niger-Congo languages, Paumari, Banawá, Mongolian, etc.	CV-languages etc.

The chart in (2) demonstrates that all the four logical possibilities (viz., allowing for initial empty c's only, word-internal hiatus only, both, or neither) are attested, although the type in which vowel-initial words are allowed but no medial onsetless syllables are tolerated seems to be more common.

7.3 Conclusion

Strict CV/VC phonology has largely concentrated on the licensing/silencing of empty nuclei: how their inherent loudness is damaged. This chapter has attempted to switch the attention to C positions, which are inherently mute, but are nevertheless dispreferred in languages when empty. It is proposed that empty onsets are not problematic in their own right, but because all these C's are followed by filled V's, endowed by governing and licensing capacities. The melodic emptiness of a C position, however, prevents the V from fully exploiting its governing charge. Therefore, the cross-linguistic markedness of onsetless syllables provides additional support for the claim that there exists some kind of communication between the melodies of adjacent phonological objects: an assumption which has also helped solve the English cross-word puzzle.

References

Abercrombie, David (1967) Elements of General Phonetics. Chicago: Aldine.

Andersen, Henning (1986) (ed.) Sandhi phenomena in the languages of Europe. Trends in Linguistics, Studies and Monographs 33. Berlin – New York – Amsterdam: Mouton de Gruyter.

Anderson, John M. (2004) Contrast in phonology, structural analogy and the interfaces. Studia Linguistica 58.3: 269-287.

Anderson, Stephen R. (1982) The analysis of French shwa: or, how to get something for nothing. Language 58.3: 534-573.

Anderson, John and Colin Ewen (1987) Principles of Dependency Phonology. Cambridge: CUP.

Aoun, Youssef (1979) Is the syllable or the supersyllable a constituent? MIT Working Papers in Linguistics 1: 140-148.

Aronoff, Mark (1976) Word Formation in Generative Grammar. Linguistic Inquiry Monograph 1. Cambridge, MA: MIT Press.

Backley, Philip (1993) Coronal: the undesirable element. UCL WPL 8: 301-323.

Blaho, Sylvia (2003) Derived environment effects in Optimality Theory: The case of pre-sonorant voicing in Slovak. Paper presented at the 11th Annual Workshop on Formal Approaches to Slavic Linguistics, University of Massachusetts, Amherst.

Blaho, Sylvia (2004) Interactions of sonorant and obstruent voicing. MA thesis, PPKE, Piliscsaba.

Blevins, Juliette (1995) The syllable in phonological theory. In Goldsmith (1995): 206-244.

Booij, Geert E. (1986) Two cases of external sandhi in French: enchaînement and liaison. In Andersen (1986): 93-103.

Booij, Geert E. (1999) The role of the prosodic word in phonotactic generalizations. In T. Alan Hall and Ursula Kleinhenz (eds.) Studies on the Phonological Word. Amsterdam/Philadelphia: John Benjamins. 47-72.

Borowsky, Toni (1986) Topics in the Lexical Phonology of English. PhD dissertation, University of Massachusetts.

Broadbent, Judith (1991) Linking and intrusive r in English. UCL Working Papers in Linguistics 3: 281-302.

Brockhaus, Wiebke (1995b) Skeletal and suprasegmental structure within Government Phonology. In Durand and Katamba (1995): 180-221.

Bromberger, Sylvain and Morris Halle (1989) Why phonology is different. Linguistic Inquiry 20.1: 51-70.

Bronstein, Arthur J. (1960) The Pronunciation of American English. New York: Appleton-Century-Crofts.

Burzio, Luigi (1994) Principles of English Stress. Cambridge: CUP.

Casali, Rod (1997) Vowel elision in hiatus contexts: which vowel goes? Language 73.3: 493-533.

Charette, Monik (1989) The Minimality Condition in phonology. Journal of Linguistics 25: 159-187.

Charette, Monik (1990) Licence to govern. Phonology 7: 233-253.

Charette, Monik (1991) Conditions on Phonological Government. Cambridge: CUP.

Charette, Monik (1992) Mongolian and Polish meet Government Licensing. SOAS Working Papers in Linguistics and Phonetics 2: 275-292.

Chomsky, Noam (1981) Lectures on Government and Binding. Dordrecht: Foris.

Chomsky, Noam and Morris Halle (1968) The Sound Pattern of English. New York: Harper and Row.

Clements, George N. and Samuel Jay Keyser (1983) CV Phonology: A Generative Theory of the Syllable. Cambridge, Mass.: The MIT Press.

Coleman, John (1995) Declarative lexical phonology. In Durand and Katamba (1995): 333-382.

Cyran, Eugeniusz (1998) (ed.) Structure and Interpretation: Studies in phonology. Lublin: Folium.

Cyran, Eugeniusz (2003) Complexity Scales and Licensing Strength in Phonology. Lublin: Kul.

Csides, Csaba (1998) Element A: representing the third English glide. Ms., ELTE, Budapest.

Csides, Csaba (2000) Aspects of consonant lenition and phonotactics. Ms., ELTE, Budapest.

Csides, Csaba (2002) License to properly govern. SOAS Working Papers in Linguistics and Phonetics 12: 67-88.

Csides, Csaba (2004) Bidirectional government in strict CV: Evidence from English. SOAS Working Papers in Linguistics and Phonetics 13: 81-126.

Davidsen-Nielsen, Niels (1974) Syllabification in English words with medial *sp, st, sk*. Journal of Phonetics 2: 15-45.

Davis, Stuart (2001) On the analysis of aspirated weak syllables in English. Lecture given at Stanford.

Davis, Stuart (2003) The footing of dactylic sequences in American English. In Takeru Homna, Masao Okazaki, Toshiyuki Tabata, and Shin-ichi Tanaka (eds.) A New Century of Phonology and Phonological Theory. Tokyo: Kaitakusha. 277-289.

Davis, Stuart (2005) "Capitalistic" vs. "militaristic": The paradigm uniformity effect reconsidered. In Laura Downing, T. A. Hall, and Renate Raffelsieffen (eds.) Paradigms in Phonological Theory. Oxford: OUP.

Dienes, Péter (2000) VC Phonology: a theory of consonant lenition and phonotactics. MA thesis, ELTE, Budapest.

Dienes, Péter and Péter Szigetvári (1999) Repartitioning the skeleton: VC Phonology. Ms., ELTE, Budapest.

Dubach Green, Antony (2003) American English "R-colored" vowels as complex nuclei. Glot International 7.5: 122-123.

Durand, Jacques (1995) Universalism in phonology: atoms, structures and derivations. In Durand and Katamba (1995): 267-288.

Durand, Jacques and Francis Katamba (1995) (eds.) Frontiers of Phonology: Atoms, Structures, Derivations. Harlow: Longman.

EPD = Searchable English Pronunciation Dictionary: http://seas3.elte.hu/epd.html.

Fleischhacker, Heidi (2001) Cluster-dependent epenthesis asymmetries. UCLA Working Papers in Linguistics 7, Papers in Phonology 5: 71-116.

Fox, Robert A. and Dale Terbeek (1977) Dental flaps, vowel duration and rule ordering in American English. Journal of Phonetics 5: 27-34.

Giegerich, Heinz J. (1992) English Phonology: An Introduction. Cambridge: CUP.

Gimson, Alfred Charles (1980) An Introduction to the Pronunciation of English. London: Edward Arnold.

Goldsmith, John A. (1976) Autosegmental Phonology. PhD dissertation, MIT. (Published by Garland Press, New York, 1979.)

Goldsmith, John A. (1995) (ed.) The Handbook of Phonological Theory. Cambridge, Mass. and Oxford: Blackwell.

Gussenhoven, Carlos (1986) English plosive allophones and ambisyllabicity. Gramma 10: 119-141.

Gussmann, Edmund and Jonathan Kaye (1993) Polish notes from a Dubrovnik café: I. The yers. SOAS Working Papers in Linguistics and Phonetics 3: 427-462.

Halle, Morris and K. P. Mohanan (1985) Segmental phonology of modern English. Linguistic Inquiry 16.1: 57-116.

Hammond, Michael (1999) The Phonology of English: A Prosodic Optimality Theoretic Approach. Oxford: OUP.

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

Harris, John (1992a) Licensing inheritance. UCL Working Papers in Linguistics 4: 359-406.

Harris, John (1992b) The autosegmental effects of prosodic licensing. Paper presented at the Government Phonology Workshop, 7th International Phonology Meeting, Krems.

Harris, John (1994) English Sound Structure. Cambridge, Mass. and Oxford: Blackwell.

Harris, John and Jonathan Kaye (1990) A tale of two cities: London glottalling and New York City tapping. The Linguistic Review 7: 251-274.

Harris, John and Geoff Lindsey (1993) There is no level of phonetic representation. UCL Working Papers in Linguistics 5: 355-373.

Harris, John and Geoff Lindsey (1995) The elements of phonological representation. In Durand and Katamba (1995): 34-79.

Hayes, Bruce (1982) Extrametricality and English stress. Linguistic Inquiry 13: 227-276.

Hayes, Bruce (1989a) Compensatory lengthening in Moraic Phonology. Linguistic Inquiry 20: 253-306.

Hayes, Bruce (1989b) The prosodic hierarchy in meter. In Paul Kiparsky and G. Youmans (eds.) (1989) Rhythm and meter. Orlando: Academic Press.

Hayes, Bruce (2000) Gradient well-formedness in Optimality Theory. In Joost Dekkers, Frank van der Leeuw and Jeroen van de Weijer (eds.) (2000) Optimality Theory: Phonology, Syntax, and Acquisition. Oxford: OUP. 88-120.

Higginbottom, Eleanor (1964) Glottal reinforcement in English. Transactions of the Philological Society, London: 129-142.

Hoard, James E. (1971) Aspiration, tenseness and syllabication in English. Language 47.1: 133-140.

Honeybone, Patrick (1999) 'I blame the government'. Language Sciences 21: 177-221.

Honeybone, Patrick (2005) Sharing makes us stronger: process inhibition and segmental structure. In Philip Carr, Jacques Durand and Colin J. Ewen (eds.) Headhood, Elements, Specification and Contrastivity: Phonological Papers in Honour of John Anderson. Amsterdam: John Benjamins.

Hooper, Joan Bybee (1976) Word frequency in lexical diffusion and the source of morphophonological change. In William M. Christie jr. (ed.) Current Progress in Historical Linguistics. Proceedings of the Second International Conference on Historical Linguistics, Tucson, Arizona, 12-16 January 1976. Amsterdam-New York-Oxford: North-Holland Publishing Company. 95-105.

Hooper, Joan Bybee (1978) Constraints on schwa-deletion in American English. In Jacek Fisiak (ed.) Recent Developments in Historical Phonology. Berlin/New York: Mouton de Gruyter. 183-207.

Hulst, Harry van der and Nancy A. Ritter (1999) Head-driven phonology. In Harry van der Hulst and Nancy A. Ritter (eds.) The Syllable: Views and Facts. Berlin/New York: Mouton de Gruyter. 113-167.

Inkelas, Sharon (1989) Prosodic Constituency in the Lexicon. PhD dissertation, Stanford University. (Published by Garland Press, New York, 1990.)

Inkelas, Sharon and Draga Zec (1990) (eds.) The phonology-syntax connection. Chicago: University of Chicago Press.

Inkelas, Sharon and Draga Zec (1995) Syntax-phonology interface. In Goldsmith (1995): 535-549.

Itô, Junko (1989) A prosodic theory of epenthesis. NLLT 7: 217-260.

Iverson, Gregory K. and Joseph C. Salmons (1995) Aspiration and laryngeal representation in Germanic. Phonology 12. 369–396.

Jensen, John T. (1987) English stop allophones in metrical theory. Proceedings of the XIth International Congress of Phonetic Sciences, Volume 2: 153-156.

Jensen, John T. (2000) Against ambisyllabicity. Phonology 17.2: 187-235.

Jones, Daniel (1960) An Outline of English Phonetics. 9th edition. Cambridge: CUP.

Jones, Daniel (1966) The Pronunciation of English. Cambridge: CUP.

Kahn, Daniel (1976) Syllable-Based Generalisations in English Phonology. PhD dissertation, MIT. (Published by Garland Press, New York, 1980.)

Kaisse, Ellen M. (1985) Connected Speech: The Interaction of Syntax and Phonology. Orlando: Academic Press.

Kálmán, László (1989) Monotonicity in phonology. Acta Linguistica Hungarica Vol. 39 (1-4): 133-147.

Kaye, Jonathan (1990a) Government in phonology. The case of Moroccan Arabic. The Linguistic Review 6: 131-159.

Kaye, Jonathan (1990b) 'Coda' licensing. Phonology 7: 301-330.

Kaye, Jonathan (1990c) Whatever happened to dialect B? In Joan Mascaró and Marina Nespor (eds.) Grammar in Progress: GLOW Essays for Henk van Riejmsdijk. Dordrecht: Foris. 259-263.

Kaye, Jonathan (1992) Do you believe in magic? The story of s+C sequences. SOAS Working Papers in Linguistics and Phonetics 2: 293-313.

Kaye, Jonathan (1995) Derivations and interfaces. In Durand and Katamba (1995): 289-332.

Kaye, Jonathan and Jean Lowenstamm (1986) Compensatory lengthening in Tiberian Hebrew. In L. Wetzels and E. Sezer (eds.) Studies in Compensatory Lengthening. Dordrecht: Foris. 97-132.

Kaye, Jonathan, Jean Lowenstamm and Jean-Roger Vergnaud (1985) The internal structure of phonological representations: a theory of charm and government. Phonology Yearbook 2: 305-328.

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

Kenstowicz, Michael (1994) Phonology in Generative Grammar. Cambridge, Mass. and Oxford: Blackwell.

Kim, Chin-Wu (1970) A theory of aspiration. Phonetica 21: 107-116.

Kiparsky, Paul (1979) Metrical structure assignment is cyclic. Linguistic Inquiry 10: 421-442.

KLV = Kaye, Jonathan, Jean Lowenstamm and Jean-Roger Vergnaud

Kreidler, Charles (1989) The Pronunciation of English. Cambridge, Mass. and Oxford: Blackwell.

Ladd, D. Robert and James M. Scobbie (2003) External sandhi as gestural overlap? Counter-evidence from Sardinian. In J. Local, R. Ogden and R. Temple (eds.) Papers in Laboratory Phonology VI. Cambridge: CUP. Chapter 9: 162-180.

Ladefoged, Peter (1993) A Course in Phonetics. 3rd edition. Orlando, FL: Harcourt Brace Jovanovich.

Ladefoged, Peter and Ian Maddieson (1996) The Sounds of the World's Languages. Cambridge, MA: Blackwell.

Larsen, Uffe Bergeton (1995) Vowel length, Raddoppiamento Sintattico and the selection of the definite article in Modern Italian. In Léa Nash, Georges Tsoulas,

Anne Zribi-Hertz (eds.). Actes du deuxiéme colloque Langues et Grammaire. Université Paris 8. 110-124.

Laver, John (1994) Principles of Phonetics. Cambridge: CUP.

Leslie, David (1983) Left capture and British voiceless stop allophony. Ms., London: University College.

Lindsey, Geoff and John Harris (1990) Phonetic interpretation in generative grammar. UCL Working Papers in Linguistics 2: 355-369.

Lowenstamm, Jean (1996) CV as the only syllable type. In Jacques Durand and Bernard Laks (eds.) Current Trends in Phonology: Models and Methods. European Studies Research Institute, University of Salford Publications: 419-442.

Lowenstamm, Jean (1999) The beginning of the word. In Rennison and Kühnhammer (1999): 153-166.

LPD = Wells, John C. (ed.) (1990) Longman Pronunciation Dictionary. London: Longman.

Majdi, B. and D. Michaels (1987) Syllable structure, gemination and length in Iraqui. Paper presented at the 62nd Annual Meeting of the Linguistic Society of America.

McCarthy, John J. (1982) Prosodic structure and expletive infixation. Language 58: 574-590.

McCarthy, John J. (1991) Synchronic rule inversion. In Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society, February 15-18. 192-207.

McCrary, Kristie (2002) Syllable structure vs. segmental phonotactics: the phonetics and phonology of segment duration in Italian. Paper presented at the 9th International Phonology Meeting, Vienna.

Nespor, Marina (1986) The phonological word in Greek and Italian. In Andersen (1986): 65-74.

Nespor, Marina and Irene Vogel (1986) Prosodic Phonology. Dordrecht: Foris.

Newson, Mark (2004) Deforestation in syntax. In László Varga (ed.) The Even Yearbook 6 (2004). ELTE SEAS Working Papers in Linguistics, Dept. of English Linguistics, SEAS, Eötvös Loránd University (ELTE), Budapest. 135-148.

Pagliano, Claudine and Olivier Rizzolo (2000) VC versus CV: the French trial. Paper presented at the Budapest Worshop on the Skeleton, 29-30 May 2000.

Pater, Joe (2000) Nonuniformity in English stress: the role of ranked and lexically specific constraints. Phonology 17.2: 237-274.

Patterson, David and Cynthia M. Connine (2001) Variant frequency in flap production. A corpus analysis of variant frequency in American English flap production. Phonetica 2001; 58: 254-275.

Pigott, Glyne L. (1999) At the right edge of words. The Linguistic Review 16-2: 143-185.

Polgárdi, Krisztina (1998) Vowel Harmony. An Account in Terms of Government and Optimality. The Hague: Holland Academic Graphics.

Prince, Alan and Paul Smolensky (1993) Optimality Theory: Constraint Interaction in Generative Grammar. (http://roa.rutgers.edu/view.php3?roa=537)

Prince, Alan and Paul Smolensky (2004) Optimality Theory: Constraint Interaction in Generative Grammar. Cambridge, Mass. and Oxford: Blackwell.

Rennison, John (1998) Contour segments without subsegmental structures. In Cyran (1998): 227-245.

Rennison, John (1999) Can there be empty phonological elements? On empty heads and empty operators. In Rennison and Kühnhammer (1999): 183-193.

Rennison, John and Klaus Kühnhammer (1999) (eds.) Phonologica 1996. Syllables!? The Hague: Holland Academic Graphics.

Ritter, Nancy A. and Robert M. Vago (1999) Subsyllabic constituency in Hungarian: Implications for moraic phonology and government phonology. In Rennison and Kühnhammer (1999): 219-245.

Rowicka, Grażyna J. (1998) The audibility and visibility of Mohawk ghosts. In Cyran (1998): 247-260.

Rowicka, Grażyna J. (1999) On Ghost Vowels. A Strict CV Approach. The Hague: Holland Academic Graphics.

Rubach, Jerzy (1996) Shortening and ambisyllabicity in English. Phonology 13. 197-237.

Rubin, Dominic (2001) The Licensing and Interpretation of Coronality: A New Approach. PhD dissertation, SOAS, London.

Rubin, Dominic (to appear) Coronality, syllable structure, and government. To appear in Phonologica 2002.

Scheer, Tobias (1996) Une théorie de l'interaction directe entre consonnes. PhD dissertation, Université Paris 7.

Scheer, Tobias (1997) Vowel-zero alternations and their support for a theory of consonantal interaction. In Pier Marco Bertinetto, Livio Gaeta, Georgi Jetchev and David Michaels (eds.). Certamen Phonologicum III. Papers from the Third Cortona Phonology Meeting, April 1996. Torino: Rosenberg and Seiler. 67-88.

Scheer, Tobias (1998a) A unified model of Proper Government. The Linguistic Review 15: 41-67.

Scheer, Tobias (1998b) Governing domains are head-final. In Cyran (1998): 261-285.

Scheer, Tobias (1999a) A theory of consonantal interaction. Folia Linguistica XXXII.3-4: 201-237.

Scheer, Tobias (1999b) Syllable unstructure: CVCV. Handout at the 6th Central European Summer School in Generative Grammar, Plovdiv, August 1999.

Scheer, Tobias (2001) A representational theory of morphological information in phonology. Paper presented at Generative Linguistics in Poland 3, Warsaw, 7-8 April 2001.

Scheer, Tobias (2004) A Lateral Theory of Phonology. Vol 1: What is CVCV, and Why Should it Be? Berlin: Mouton de Gruyter.

Scheer, Tobias (to appear) A Lateral Theory of Phonology. Vol 2: On Locality, Morphology and Phonology in Phonology. Berlin: Mouton de Gruyter.

Ségéral, Philippe and Tobias Scheer (1999a) The Coda Mirror. Ms., Université de Paris 7 and Université de Nice.

Ségéral, Philippe and Tobias Scheer (1999b) Full interpretability: evidence from virtual geminates. Paper presented at the 32nd Poznan Linguistic Meeting, 30 April - 2 May 1999.

Ségéral, Philippe and Tobias Scheer (1999c) Is the Coda Mirror a phonological object? Paper presented at the First Annual Meeting of the GDR 1954 "Phonologie" on Lenition and Fortition, Nice 24-25 June 1999.

Seigneur-Froli, Delphine (2004) Diachronic consonant lenition and exotic word-initial clusters in Greek: a unified account. Ms., Université de Nice.

Seigneur-Froli, Delphine (to appear) Greek *pteron*: neither extrasyllabicity nor magic. To appear in Phonologica 2002.

Selkirk, Elizabeth O. (1972) The Phrase Phonology of English and French. MIT dissertation. (Published by Garland Press, New York, 1980.)

Selkirk, Elizabeth O. (1980) Prosodic domains in phonology: Sanskrit revisited. In Mark Aronoff and M. L. Kean (eds.) Juncture. Saratoga, Calif.: Anma Libri. 107-129.

Selkirk, Elizabeth O. (1982, written in 1978) The syllable. In Harry van der Hulst and Norval Smith (eds.) The Structure of Phonological Representations, Part II. Dordrecht: Foris Publications. 337-383.

Selkirk, Elizabeth O. (1986) On derived domains in sentence phonology. Phonology 3: 371-405.

Smith, Jennifer L. (2002) Phonological Augmentation in Prominent Positions. PhD dissertation, University of Massachusetts, Amherst. (Published by Routledge, 2004.)

SPE = Chomsky and Halle (1968)

Spencer, Andrew (1996) Phonology: Theory and Description. Cambridge, Mass. and Oxford: Blackwell.

Stampe (1972) How I Spent My Summer Vacation. PhD dissertation, Ohio State University.

Steriade, Donca (1999) Alternatives to syllable-based accounts of consonantal phonotactics. Ms.

Steriade, Donca (2000) Paradigm uniformity and the phonetics-phonology boundary. In J. Pierrehumbert and M. Broe (eds.) Papers in Laboratory Phonology. Vol.5. Cambridge: CUP. 313-334.

Szigetvári, Péter (1999a) Against word-final empty nuclei. Poster presented at the 7th Manchester Phonology Meeting (14 May 1999).

Szigetvári, Péter (1999b) VC Phonology: A Theory of Consonant Lenition and Phonotactics. PhD dissertation, MTA/ELTE, Budapest.

Szigetvári, Péter (2000) Why CVCV. In László Varga (ed.) The Even Yearbook 4 (2000). ELTE SEAS Working Papers in Linguistics, Dept. of English Linguistics, SEAS, Eötvös Loránd University (ELTE), Budapest. 117-152.

Szigetvári, Péter (2002a) Branching onsets and syncope in English. Paper presented at the Conference on English Phonology, Université de Toulouse, 26-28 June 2002 and at the BuPhoC session of 2 October 2002.

Szigetvári, Péter (2002b) Syncope in English. In László Varga (ed.) The Even Yearbook 5 (2002). ELTE SEAS Working Papers in Linguistics, Dept. of English Linguistics, SEAS, Eötvös Loránd University (ELTE), Budapest. 139-149.

Szigetvári, Péter (2003) The null-hypotheses of syllable structure. Lecture held at PLM34, NP – GP Workshop, Poznań, 1-3 May 2003.

Szigetvári, Péter (to appear) Branching onsets and syncope in English. Language Sciences.

Takahashi, Toyomi (1993) A farewell to constituency. UCL Working Papers in Linguistics 5: 375-410.

Ternes, Elmar (1986) A grammatical hierarchy of joining. In Andersen (1986): 11-21.

Toft, Zoe (2002) The asymmetrical behaviour of syllabic sonorants in Southern British English. In Marjo van Koppen, Joanna Sio, Mark de Vos (eds.) Proceedings of Console X. 215-229.

Trousdale, Graeme (2002) Variable ambisyllabicity. English Language and Linguistics 6.2: 267-282.

Van Dam, Mark and Paul Weaver (2001) Aspiration of stressless intervocalic voiceless stops in English. Ms., Indiana University, Bloomington.

van Oostendorp, Marc (2000) Phonological Projection. A Theory of Feature Content and Prosodic Structure. Berlin and New York: Mouton de Gruyter.

Vaux, Bert (2002) Aspiration in English. Ms., Harvard University.

Vogel, Irene (1986) External sandhi rules operating between sentences. In Andersen (1986): 55-64.

Watson, Janet C. E. (2002) The Phonology and Morphology of Arabic. Oxford: OUP.

Wells, John C. (1982) Accents of English. Cambridge: CUP.

Wiese, Richard (1996) The Phonology of German. Oxford: OUP.

Withgott, Mary Margaret (1982) Segmental Evidence for Phonological Constituents. PhD dissertation, University of Texas, Austin.

Yoshida, Shohei (1993) Licensing of empty nuclei: The case of Palestinian vowel harmony. The Linguistic Review 10: 127-159.

Zec, Draga (1988) Sonority Constraints on Prosodic Structure. PhD dissertation, Stanford University.

Zec, Draga and Sharon Inkelas (1990) Prosodically constrained syntax. In Inkelas and Zec (1990): 365-378.

Zwicky, Arnold M. (1972) Note on a phonological hierarchy in English. In Robert P. Stockwell and Ronald K. S. Macaulay (eds.) Linguistic Change and Generative Theory. Bloomington and London: Indiana University Press. 275-301.