vocoids and their prosodic distribution, with special reference to Italian and Arabic

Bosisio, Nicole

VOCOIDS AND THEIR PROSODIC DISTRIBUTION, WITH SPECIAL REFERENCE TO ITALIAN AND ARABIC

by

Nicole Bosisio

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

University of Durham

Department of Linguistics and English Language

2003

ABSTRACT OF THESIS

By

Nicole Bosisio

Vocoids and their Prosodic Distribution, with Special Reference to Italian and Arabic

For the degree of Ph.D., 2003

This study attempts to characterize vocoids, i.e. vowels and semivowels, as a unified class of segments. In order to do so, it investigates the main phenomena concerning the quantitative distribution of these sounds, namely syllabic alternation, length alternations, deletion and insertion. Such phenomena are best analyzed by making reference to prosodic structure, and syllable structure in particular. Therefore, both frameworks adopted in this thesis take into consideration this type of representation.

The main approach, which I refer to generally as Derivational Theory (DT), is based on the notion that surface phonetic forms are derived from underlying forms through a series of structural changes taking place at different levels of representation. This model is contrasted with the recently introduced (Prince and Smolensky 1993) Optimality Theory (OT), an output-oriented paradigm based on the parallel evaluation of candidate forms by means of universal but violable constraints. This thesis shows that OT offers some valuable insights into the phenomena under analysis, although there are areas in which it requires integration with derivational tools.

This study also makes specific reference to two languages: Ammani Arabic and Standard Italian. These diverge in their treatment of vocoids, but clear general trends may be detected which have also been found in other languages.

Statement of Copyright

ACKNOWLEDGEMENTS

First and foremost, I must thank my supervisor S.J. Hannahs for his support and guidance, and especially for never losing faith in my capabilities.

I also wish to thank everybody at the department of Linguistics and English Language of Durham University for giving me so much both from an academic and a human point of view. I learnt more in the five years I spent in Durham than in the rest of my life. The decision to close down such a matchless department is without any doubt a serious mistake, whose consequences will reverberate on the future of the whole university.

I am grateful to both the department and the Art and Humanities Research Board for providing vital financial support.

Special thanks go to Prof. Joe Emonds, without whose help and advice I would not have embarked on this journey.

I also want to take the opportunity to express my gratitude to the city of Durham for its breathtaking beauty and the friendliness of its inhabitants, which provided inspiration and serenity throughout the period of my study. I will never forget the years I spent here.

My parents, who have consistently given me help and support, occupy an important place in these acknowledgements.

I must also thank Layla, Taghreed and Mohammed for acting as unofficial victims of my experiments with Arabic. There are another thousand reasons for which I should thank Mohammed, and I will never find enough words to mention them all. Although, in his own words, he 'never quite understood what phonology is all about', he contributed more than anyone else to the accomplishment of this thesis.

To Mohammed, to my parents and my cats. And to all those who died and that I loved so desperately. And to all those who died and that I did not even get to know.

TABLE OF CONTENTS

INTRODUCTION

The main aim of this thesis is to provide a unified account of a class of segments that comprises vowels and semivowels, or glides. Following Pike (1943), I have given these sounds the collective name of *vocoids*. Vocoids all have in common a high degree of sonority, and have been traditionally characterized by means of the feature [-consonantal].

In order to investigate the behaviour of this class of segments, this study takes into consideration their prosodic distribution, by analyzing the phenomena of syllabic alternation, length alternation, deletion and insertion. These are all phonological processes that involve quantitative, rather than qualitative, changes, and are therefore best analyzed by making reference to suprasegmental constituency. More specifically, I will show that syllable structure - and, to a lesser extent, foot structure - has a fundamental role in determining the organization and distribution of vocoids.

The main theoretical framework within which the present analysis is couched is a serial, rule-based approach that I refer to as Derivational Theory (DT), in that it derives surface forms from underlying forms by means of a series of phonological operations belonging to different levels of representation. This model comprises several subtheories which deal with the different aspects of phonological representations, such as Autosegmental Phonology, Prosodic Phonology, Lexical Phonology, Moraic Theory or Templatic Syllabification. Such an approach will be contrasted with Optimality Theory (OT), a more recent model based on the parallel evaluation of output candidates by means of violable universal constraints. A second aim of this thesis is therefore to find out whether the above-mentioned prosodic phenomena can be characterized more effectively within one or the other framework.

In order to constrain the analysis within the boundaries of a Ph.D. thesis, this study refers specifically to two languages: Italian and Arabic. Since these are not genetically related (if not in a remote, debatable way), a comparison between them may offer particularly interesting insights into the issues in hand. Both languages comprise a considerable number of local varieties, all differing in their phonological behaviour. Therefore, a specific variety has been selected for each language, though reference to other dialects will be made throughout the thesis. The choice was easier in the case of Italian, which has a clearly defined standard used in all social circumstances by a reasonably large part of the native population. This variety has been simply called

Standard Italian (SI) or, for short, *Italian*. In the Arab countries, on the other hand, what is generally labelled as (Modern) Standard Arabic is a diastratically and diaphasically marked variety which is rarely used in everyday speech. Therefore, I have chosen to analyze the dialect that is most accessible to me, that is an urban variety spoken by the educated population of Amman, the capital of Jordan, which I have termed *Ammani Arabic* (AA). In his socio-linguistic account of Jordanian Arabic, Suleiman (1985) distinguishes at least three sub-varieties of this dialect: the *Madani*, or urban, one, the *Fallahi* one, spoken in the countryside by the sedentary population, and the *Bedouin* one. AA corresponds roughly to the Madani variety as it is spoken in Amman.

This study is divided into five chapters, plus an introduction and a conclusion.
The first chapter provides an overview of the theoretical background that will be adopted throughout the rest of the thesis. The initial section introduces the category of vocoids, giving a preliminary definition of this class of sounds, first from a phonetic and then from a phonological perspective. It then argues, anticipating the discussion in chapter 2, that glides and vowels may be classified as part of the same category on the grounds that homorganic glides and vowels share the same set of features, and are thus generally undistinguishable underlyingly.
The following three sections provide the tools for an analysis of suprasegmental structure within a DT framework. Section 1.2 first introduces Autosegmental Phonology and the different models of the skeletal tier, then offers an overview of Prosodic Phonology and of Lexical Phonology. Finally, it discusses the non-concatenative morphology of Arabic. Section 1.3 is entirely dedicated to syllable structure, which is the most important trigger for phonological alternations of the type analyzed in this thesis. It first looks at the internal make-up of syllables, then at syllable structure constraints, at the mora as a unit of syllabic weight and finally at the process of syllabification. The next section deals with stress assignment. After giving a definition of prominence, it introduces representational models such as prosodic trees and metrical grids. It then goes on to describe foot structure, which is intimately related to stress location, word stress and the relationship between stress assignment and syllable structure.
The final section provides an outline of OT, first explaining the basic notions of universality and violability, constraint ranking, constraint conflict and candidate

evaluation. It then goes on to introduce different types of constraints, concentrating especially on prosodic and morpho-phonological structure, concluding with a brief account of output-to-output correspondence.

Each of the remaining four chapters describes a specific phenomenon and includes three sections: one giving a general account of the processes under analysis and a review of the relevant literature, and two dealing with the occurrences of such phenomena in Italian and Arabic, respectively. An alternative OT analysis is given for each of the illustrated processes.

In the second chapter, I give an account of vowel ~ glide alternations, and of the way in which these are articulated with respect to syllabic structure. The fact that homorganic glides and vowels can generally be derived from the same underlying segments, so that the difference between them can be ascribed entirely to syllabification, is the clearest substantiation of the claim that vocoids can be analyzed as a unified class. However, it will be shown that glides and vowels may, in some specific instances, be accorded independent phonological status, i.e. that there are vocoids that always surface as syllabic or non-syllabic independently of the environment. A related issue is the question of what segments should be classified as glides, and hence as vocoids.

The third chapter tackles the issue of vocoid length and length alternations. Although touching briefly on glide gemination, it looks especially at the way in which vowels are lengthened or shortened according to the prosodic environment. It first discusses the methods by which quantity can be represented, and draws a distinction between underlying and phonetic length. It then goes on to show how length alternations are generally related to syllable and foot structure constraints, and to the distribution of stress.

In the fourth chapter, I illustrate the phenomenon of vocoid deletion in its various forms. Once again, priority is given to processes affecting vowels, which can be seen as the prototypical vocoids. Deletion is portrayed as a reduction phenomenon, and various ways of representing it are provided. There are three main types of Vowel Deletion: elision, syncope and truncation. The first process is a repair strategy eliminating vocoids in order to resolve clashes between syllable nuclei, while the second one removes subminimal syllables by deleting their nuclear vowels. Truncation, on the other hand, seems to be triggered by foot structure rather than syllable structure constraints, and consists of the deletion of domain-final, post-consonantal vocoids.

The fifth chapter describes the mirror-image process to deletion, i.e. insertion, or epenthesis. This is a phenomenon by which a segment which is not present in the underlying representation is phonetically realized, and is strictly related to prosodic structure requirements. The three sections first deal with Glide Insertion, by which a homorganic glide is inserted after a high vowel in order to separate it from another nuclear element, thus resolving a vowel clash. This process can actually be construed as a type of lengthening, since the features of the high vowel are copied to fill in the following onset. Vowel epenthesis is then analyzed. Generally, this is also a syllable-driven process, since it applies to insert a nuclear element with which stray consonants can be syllabified. Various aspects of Vowel Insertion are taken into consideration, such as the quality of the epenthetic vowel, the location of the epenthetic sites, the relationship with stress and the interaction with Vowel Deletion.

A sixth chapter provides conclusive remarks and further suggestions concerning the issues under analysis. It offers a general overview of vocoids and their prosodic distribution, concentrating on their relationship with syllable and foot structure and stress assignment. It then compares Italian and Arabic in their treatment of this class of segments, pointing out both similarities and differences. It finally gives a general evaluation of OT as set against DT, illustrating advantages and disadvantages of both frameworks.

METHODS OF TRANSCRIPTION

Throughout the thesis, examples are reported by means of two types of transcription: a phonetic transcription, and a transcription following the spelling practices of the languages involved. The former is used when rendering the actual pronunciation or the Underlying Representation of a word or sentence, and follows the conventions of the International Phonetic Alphabet. Since the exact quality of segments is rarely relevant to the phenomena analyzed in this thesis, phonetic transcription is generally rather broad. Stress marks are only transcribed when relevant to the specific argument, with ['] representing main word stress, [,] secondary stress.

As for the 'spelling' transcription, it is often adopted when there is a close correspondence between spelling and pronunciation, or when the segments whose transcription is not transparent are not relevant to the discussion in hand. For languages which make use of Roman characters, this type of transcription corresponds to their spelling conventions, except when the location of stress is signalled. This is done, for all words, by means of a stress mark on the relevant vowel, which is acute when representing main word stress (e.g. *á*), grave when representing secondary stress (e.g. *à*). This does not always correspond to the spelling conventions of languages like Italian, which often report main stress by means of a grave mark.

Among those linguistic varieties which adopt a different script, a special transliteration system is used to report examples from Arabic. Examples from other languages which do not use Roman characters, such as Greek or Japanese, will be reported in whatever form was found in the source.

The transliteration system that is used to convert Arabic letters into Roman characters, and which should not be confused with phonetic transcription, is shown in tables (1) – (3) below.

Table (1): consonants

ARABIC LETTERS	ROMAN LETTERS	ARABIC LETTERS	ROMAN LETTERS
ء	'	ض	D
ب	b	ط	T
ت	t	ظ	DH, Z

ث	th	ع	9
ج	j	غ	gh
ح	H	ف	f
خ	x	ق	q
د	d	ك	k
ذ	dh	ل	l
ر	r	م	m
ز	z	ن	n
س	s	ه	h
ش	sh	و	w
ص	S	ي	y
ة	a		

Table (2): vowels

ــــَــــ	a
ــــُــــ	u
ــــِــــ	i
ا	aa, a
و	uu, u
ي	ii, i
ى	a

Table (3): diphthongs

... و	au
... ي	ai

CHAPTER 1: AN OVERVIEW

1.1 Vocoids

This first section aims at giving a tentative definition of vocoids as a class of sounds, first looking at their phonetic, then at their phonological characteristics. Anticipating the discussion in chapter 2, it will be argued that, in spite of their difference in syllabicity, glides and vowels may be grouped in the same category, namely that of vocoids, based on the fact that the feature content of glides like [j] and [w] is analogous to that of the homorganic vowels [i] and [u]. Finally, subsections 1.1.2 and 1.1.3 will give a brief description of the vocoids of Italian and Arabic respectively.

1.1.1 Vocoids as a class

1.1.1.1 A phonetic definition

In this thesis, the term *vocoids* is used to define a phonological category, including *vowels* and *glides* (or *semivowels* and *semiconsonants*, see below). In most studies, however, the term has been used in a purely phonetic sense, to define the way vowels and glides are produced as opposed to consonants (or *contoids*). In fact, Pike (1943), who first introduced the distinction between vocoids and contoids, defined the two categories on the basis of their phonetic characteristics, as follows:

> A *vocoid* is a sound during which air leaves the mouth over the center of the tongue (nonlateral) and without friction in the mouth (but friction elsewhere does not affect the classification). A *contoid* is anything else (any nonvocoid).
>
> (ibid.:78)

He later (p. 143) adds that vocoids are the sounds that 'function most frequently as syllabics', and that they include practically all 'vowels' – except for 'fricative vowels' – and 'vowel glides' such as the English [w], [j] and [ɹ]. The distinction is based on the fact that vocoids have a less radical degree of constriction than contoids. More specifically, vocoids are sounds that are produced with open approximation –which means that the restriction imposed on the air stream is not sufficient to produce friction (Donegan 1985:45) – and that characteristically form the nucleus of a syllable. From this perspective, therefore, a vowel can simply be defined as a syllabic vocoid. From an acoustic point of view, vocoids are characterized by a high degree of periodicity, which gives them their 'sounding' quality (Donegan 1985:46), making them the closest sounds to song. In this respect, it has been noted (Lenneberg 1970:6) that vowel-like cooings are the first sounds produced by infants in their sixth to eighth week of life. All of these characteristics correspond to a high degree of sonority, and can be the basis for a phonological definition of vocoids. Among vocoids, however, there are clearly differences in sonority, which are reflected in differences in height, since a higher tongue position produces a greater constriction in the vocal tract.[1] Low vowels like [a] have in fact the highest degree of openness and periodicity (Angoujard 1990:14), which is why they always constitute a syllable peak. High vowels like [u] and [i], on the other hand, are less sonorous, and thus more prone to occupy 'non-nuclear' (cf. 1.3 for discussion) positions.[2] In the latter case, they are labelled as glides, and represented as [w] and [j] respectively (see *glide formation*, in 2.1). Mid vowels like [e] and [o] are somewhere in between, and whether or not they can alternate with glides is debatable. More sonorous vocoids, having a higher first formant, also exhibit a higher degree of intrinsic intensity. Apparently, all languages have at least two different vowel heights, even at a phonological level (cf. Donegan 1985:50).

When dealing with issues like sonority, vowel height and especially syllable structure, we are clearly bordering on the field of phonology, which I will soon be looking into in more detail. As far as phonetics is concerned, it must be added that scholars have often disagreed on the classification of glides. It has often been noted that their characteristics are actually half-way between those of vowels and consonants, as regards both their degree of stricture and their formant patterns, which vary appreciably from those of high vowels (see for example Maddieson and Emmorey 1985:171). In fact, authors such as Zwanenburg (1966:32), Lass (1976:17-8) and Di Luzio (1988:303) classify glides as consonants, while other authors, such as Rea (1979:355) and Gimson (1989:33-4), consider them as vowel-like segments. Maddieson (1984:91) describes glides as '*vocoid* approximants', a special case of 'approximant *consonants*', thus showing how difficult it is to actually draw a line between vowel-like and consonant-like sounds.

---

[1] Donegan (1985:55), however, points out that the degree of sonority of a vowel cannot always be traced back to its height.

1.1.1.2 A phonological perspective

If the issue may well remain open to different interpretations at the phonetic level, it can be useful to try to draw that line as a phonological distinction, in order to be able to characterize a number of phonological processes that we will be looking at in more detail in the next chapters. In its turn, the analysis of such processes is likely to give us a further clue in the arduous task of classifying and representing vocoids.

1.1.1.2.1 Distinctive features

As in phonetics and even more so, in phonology vowels have often been set apart from all other segments, including glides, which have been classified as 'consonants'. The basis for this distinction is the role segments play in syllable and word structure, prosodic elements which are more readily dealt with in phonological than in phonetic terms. At the segmental level, however, the representation of phones and phonemes must be grounded in phonetic characteristics such as degree of constriction, height and sonority, from which distinctive features are derived.

In early generative phonology (cf. Chomsky and Halle 1968), where every segment was defined as a feature matrix without ever referring to metrical structure, a vowel was characterized by the major class features [+syllabic, -consonantal], with glides being [-syllabic, -consonantal] and all other sounds being [+consonantal]. Let us leave aside, for the time being, the question of glides and of the feature [syllabic], to which we will come back later on in this discussion, and let us concentrate on the feature that distinguishes vocoids as a class, i.e. [-consonantal]. For a segment to be [-consonantal] means, in phonetic terms, that it has no central obstruction in the oral tract. The feature [consonantal] is still largely used, and was adopted both in autosegmental phonology (see 1.2) and in feature geometry.[3]

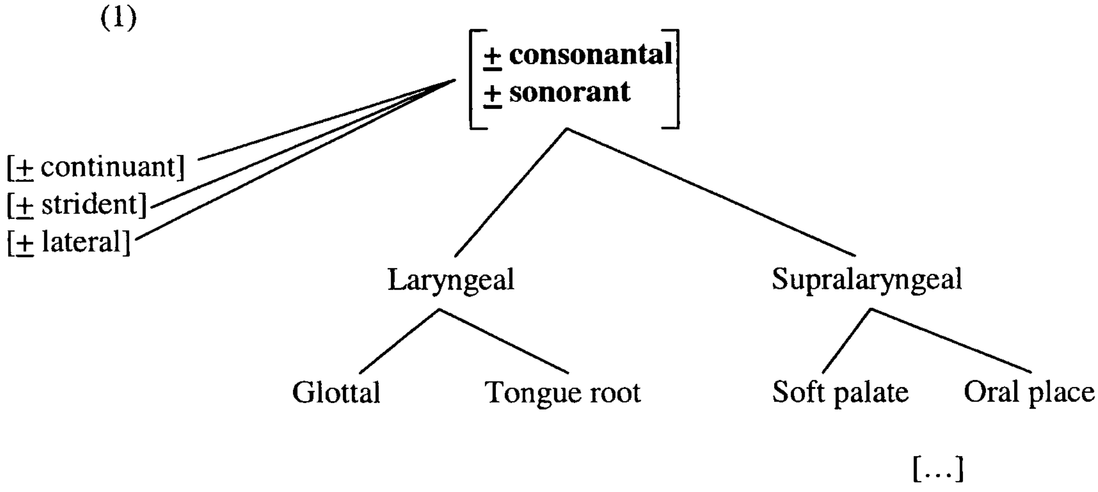
In most feature geometry models (McCarthy 1988, Halle 1992), [consonantal] is one of the root features, together with [sonorant] (cf. (1) below).

---

[2] More generally, high, less sonorous vowels are less sustainable than more sonorous vowels, and thus more susceptible to weakening processes, such as deletion (see ch. 4) or shortening (see ch. 3).

[3] Since the qualitative characterization of segments is not directly relevant to the phenomena that will be dealt with, feature geometry representations will be assumed but not explicitly portrayed throughout the thesis.

(1)



(Adapted from Halle 1992)

Both glides and vowels are [-consonantal, +sonorant], and the difference between them is a function of their position in syllable structure, i.e. of whether or not they are dominated by a syllable nucleus. There are no [-consonantal, -sonorant] sounds, which shows that a segment, in order to be [-sonorant], has to be a contoid or, in other words, that all vocoids are [+sonorant] (for the problem of voiceless vowels, see below). They are in fact characterized by a high degree of sonority, which is why they typically occupy the syllable nucleus. As was mentioned above, physical correlates of sonority are openness (corresponding to a low degree of constriction), intensity and periodicity, and vowels can be more or less sonorous according to their height. In fact, sonority is not an on-off feature, but a scalar element.

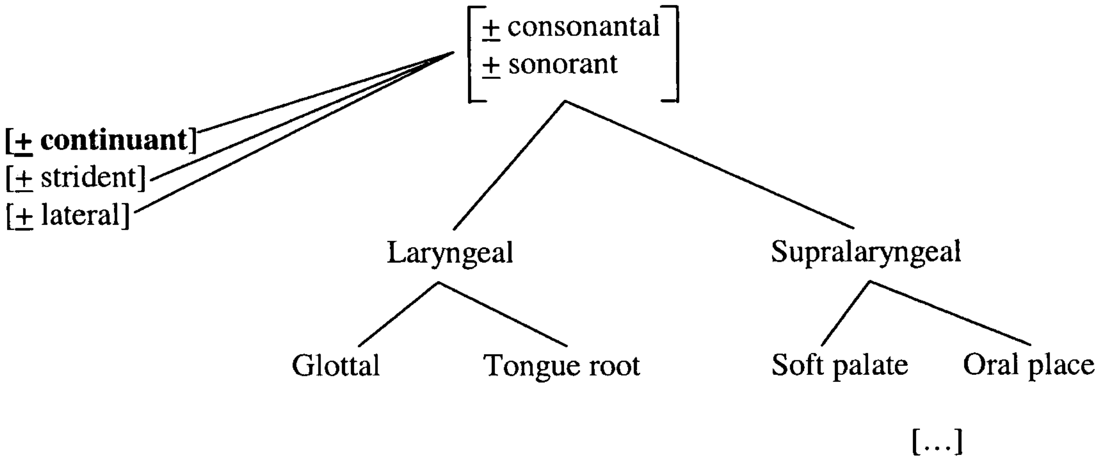
On the basis of these considerations, Angoujard (1990) claims that segments can be characterized in relation to metrical structure by referring to their degree of sonority. Based on previous works such as Kouloughli's (1978) and especially Selkirk's (1984b), he (ibid.:15) builds a sonority hierarchy for the phonemic inventory of Arabic. In such a hierarchy, segments are assigned a sonority index, with the most sonorous sound ([a]) being assigned 10 and the least sonorous (voiceless stops) being assigned 0.5. As far as other vocoids are concerned, mid vowels are assigned 9 and high vowels, together with the corresponding glides, are assigned 8. The author thus claims that so called 'major class features' such as [sonorant] and [consonantal] can be eliminated. In fact, a notion like that of 'consonantality' appears to be somewhat *ad* *hoc* and it would definitely be convenient if one could build it into an independently motivated sonority scale. However, the implementation of a sonority-based model such as Angoujard's (ibid.) is problematic. The main problem is that segments which are assigned the same index actually display different degrees of sonority, not only at a phonetic, but also at a phonological level. Most relevant here is the fact that glides are in reality less sonorous than high vowels (see above), which is why they do not constitute syllable peaks.[4] Since glides and high vowels can actually appear in the same syllable, this distinction becomes crucial to the representation of metrical structure. In a word like *wheel* [wi:l], for example, the assignment of the same index to the first and second segment obscures the fact that the degree of sonority is actually increasing slightly from the onset to the nucleus. Moreover, the phonemic inventory varies from one language to another, which may require drawing language-specific distinctions, especially since sonority differences are also a function of syllable structure. For instance, a language like Classical Arabic, which has only a two-way vowel height distinction underlyingly, will require fewer sonority indices in the higher section of the scale than English does. These considerations, however, do not undermine the fact that vocoids as a class are more sonorous than contoids, and that sonority is one of the fundamental elements in defining such classes.

Another distinctive feature that has been widely applied in generative phonology is [continuant]. It has been used mainly to distinguish [-continuant] plosives from [+continuant] fricatives, but it can be applied to the other classes of segments in that it has a clear physical correlate: the presence or absence of an obstruction in the oral tract. Therefore, this feature is also obviously related to the degree of constriction that characterizes a certain sound, and, while contoids can be either [+continuant] or [-continuant], vocoids are all [+continuant]. In spite of its fundamentally distinctive role and of its connection with sonority, this feature has not found a satisfactory representation within the framework of feature geometry; it has often been represented as hanging off the root in an awkward fashion (see (2) below), even if it has been seen by some (cf. Padgett 1991) as a dependent of Place (Kenstowicz 1994:483-4).

---

[4] However, by adopting the opposite perspective, one could also say that, when a vocoid occupies a

(2) [repeating (1)]



(Adapted from Halle 1992)

Padgett's (1991) proposal, however, would not solve the problem of how to represent the relation between continuance and sonority. This is actually quite complex, in that there are sonorant segments – namely the nasals – which have been defined as [-continuant], in that they are characterized by an obstruction in the oral cavity. The issue, however, is not of concern here: vocoids are all [+sonorant], [-consonantal] and [+continuant]. Moreover, they are typically voiced (Donegan (1985:47) describes vowels as 'the optimal voiced segments') - even if voiceless vowels can occur in many languages at a phonetic level[5] - so that they can be considered as underspecified for the feature [voice]. In fact, as is the case with all sonorant sounds, the configuration of the vocal tract cavity during the production of vocoids promotes spontaneous voicing (Lass 1976:148, Kenstowicz and Kisseberth 1979:244). As a result, if anything in a syllable is voiced, the syllable peak, i.e. the most sonorous element, is.

In conclusion, there is a clear correlation between sonority and other features, such as height, voicing, syllabicity, continuance and sustainability (Donegan 1985:55).

---

non-nuclear position, it becomes (phonetically) less sonorous.

[5] Ladefoged and Maddieson (1996:315), however, report that in some (especially Amerindian) languages voiceless vowels appear to be underlying.

1.1.1.2.2 Metrical structure

For a phonological definition to be complete, the notion of metrical structure must also be taken into consideration. From this point of view, glides differ sharply from the corresponding high vowels with which, on the other hand, they seem to share an entire set of features. In fact, [j] and [w] have traditionally been considered shorter, non-syllabic versions of [i] and [u]. As was mentioned above, in Chomsky and Halle's 'Sound Pattern of English' (1968) glides and vowels are distinguished by means of the feature [syllabic], but with the (re)introduction of the syllable as a phonological unit (cf. Fudge 1969, Kahn 1980) such a feature, in itself problematic, has become superfluous, since syllabicity can be represented as a function of the position of a segment with respect to the syllable nucleus[6] (cf. 1.3). Glides are therefore assigned the same feature structure as the corresponding high vowels, but differ from them in that they do not occupy the syllable nucleus.

We have seen that glides have often been described as consonant-like sounds at the phonetic level and that, based on their non-syllabicity, they have been classified as consonants at the phonological level, too. However, syllabicity is a dubious criterion for classifying segments into vocoids and contoids. It must be pointed out that, in many languages, consonants can function as syllable peaks: one would thus be obliged to draw an unlikely distinction between, for example, a syllabic /r/ and a non-syllabic /r/, perhaps defining the former as a 'vowel'. The distinction between vocoids and contoids must be drawn at the segmental level, while syllabic structure is best represented on a different tier and comes into play only when segments are syllabified (cf. 1.3). On the segmental tier, glides like [w] and [j] have been shown to share their distinctive features with vowels like [u] and [i], with which in many cases they alternate just as syllabic and non-syllabic consonants do. For this reason, throughout this thesis the high front vocoid, having [i] and [j] as its allophones, will be represented underlyingly as an underspecified /I/; the high back vocoid phoneme will be represented as /U/, with [u] and [w] as its allophones. As we will see in chapter 2, the phenomena of glide- and vowel-formation (or better, of glide/vowel alternation) constitute perhaps the strongest motivation for setting up a category of vocoids: the

[6] Pike (1943:78) and Zwanenburg (1966:32), however, had already pointed out that the only clear distinction could be traced on the basis of the concept of syllable.

same segment can surface either as a vowel or as a glide, according to its position with regard to other vocoids, i.e. according to its position within the syllable.

In this respect, a further distinction must be drawn between *on-glides* (or *semiconsonants*), which precede the nuclear vowel (as in *we*), and *off-glides* (or *semivowels*), which follow it (as in *pay*). The two classes of sounds differ in their phonetic make-up and even in their phonological behaviour. Off-glides are much less 'consonant-like' than on-glides; they form a *diphthong* with the preceding syllabic vocoid, and the whole cluster 'acts like one of the long, simple vowels' (O'Connor 1971:107). A diphthong can in fact be defined as the sequence of two vowels belonging to the same syllable, and while traditionally a distinction has been drawn between 'rising diphthongs' – formed by an on-glide plus a vowel – and 'falling diphthongs' – formed by a vowel plus an off-glide – it has been argued that only the latter clusters can be labelled as 'diphthongs proper' (cf. for example O'Connor 1971:107-10). Similar considerations led other scholars (e.g. Lass 1976:13-20; Serianni 1989:19-22) to conclude that only semiconsonants are 'real' glides, while semivowels are simply the second element of a diphthong.

Whatever position we may take in these issues, it is useful to introduce a three-way alternation between vowels, on-glides and off-glides, since semivowels sometimes pattern in different ways compared to semiconsonants (cf. Lass 1976:17-8).[7] Such a distinction, however, can be drawn only when looking at syllable structure, since at the segmental level high vowels such as [i] and [u] and glides such as [j] and [w] share the same distinctive features, and can generally (see 2.1) be analyzed as having identical Underlying Representations (URs). This is probably a sufficient reason to introduce a category of vocoids, including syllabic and non-syllabic [-consonantal] segments. It can be argued, however, that some glides, such as French and Spanish /j/ when spelt ⟨ll⟩, behave exactly like consonants: they do not alternate with vowels, and they show a strong tendency to become fricatives (cf. Spence 1971; Rea 1979). On the other hand, as will be shown in 2.3, a Semitic language like Arabic exhibits glides that have traditionally been classified as root consonants (cf. 1.2), even though they can actually alternate with high vowels. Hence, the question arises of whether it may be

---

[7] In reality, a four-way distinction can be traced, since so-called 'off-glides' may appear either in the nucleus or in the coda of a syllable (see 2.1 for discussion).

insightful to trace a distinction – at least at the phonological level - between 'vocalic' and 'consonantal' glides, classifying the former as vocoids, the latter as contoids.

1.1.2 Vocoids in Italian

In Italian there are seven vowel phonemes - /a/, /ɛ/, /e/, /ɔ/, /o/, /i/, /u/ - and two glides, [w] and [j]. [ɛ] and [ɔ], however, only appear in stressed position, and are often found in complementary distribution with [e] and [o] respectively. In non-standard varieties of Italian, the contrast is almost always neutralized, so that the distinctive function of /ɛ/ and /ɔ/ with respect to /e/ and /o/ is limited to a few minimal pairs such as [attʃetta] 'hatchet' ~ [attʃɛtta] 's/he accepts'. Because of its phonotactics allowing onsetless syllables and disfavouring codas (see 1.3), Italian words are rich in vocoids. These can be joined freely into diphthongs and hiati, allowing all possible combinations.

1.1.3 Vocoids in Arabic

The vocalic inventory of Arabic varies greatly from one dialect to another. In Ammani Arabic (henceforth AA), as in Standard Arabic (henceforth SA), there are six vowel phonemes, three short (/a/, /u/ and /i/), and three long (/a:/, /u:/ and /i:/) (see 3.3), but each of them has a wealth of allophones, sometimes overlapping with each other. For instance, AA /a/ tends to be pronounced as [ɑ], or even as [ɒ], in the vicinity of emphatic consonants (e.g. *D[a]ll* 'he continued', *b[a]TT* 'ducks'), as [a] in the presence of pharyngeals and velars (e.g. *9[a]sha* 'dinner', *sh[a]xs* 'person'), as [e] word-finally when preceded by a front consonant (e.g. *xidm[e]* 'service') and as [æ] in the remaining environments. The phoneme /i/, on the other hand, is often realized as [ɨ] (or even as [e]) when in contiguity with pharyngeal or emphatic consonants (e.g. *D[ɨ]ffa* '(West) Bank', *H[e]lw* 'sweet') and as [ɪ] in all other cases, while /u/ surfaces as [o] (or [ʊ]) in the environment of emphatic, pharyngeal and sometimes velar consonants (e.g. *D[o]hur* 'noon', *9[o]mur* 'age, life', *'[o]xt* 'sister'), as [u] elsewhere. The long vowels, too, exhibit various allophones, which often have a similar distribution to those of the homorganic short vowels. The picture is further complicated by the fact that some of these sounds appear in free variation in certain environments.[8]

In line with these observations, Cantineau (1960:147) remarks that, in the vicinity of emphatic consonants, the vowel /a/ tends to be backed and the vowels /i/ and /u/ tend to open into [e] and [o] respectively; the pharyngeal and laryngeal consonants, on the other hand, tend to make vowels 'sound like [a]' (i.e. to open /i/ and /u/ even more), while labial consonants tend to make them round.

Although it has been claimed that in Arabic glottals ([ʔ], [h]), and even pharyngeals ([ʕ], [ħ]), behave as glides in certain circumstances (cf. McCarthy 1981, Watson 1989), only [j] and [w] can be really considered as such (see Angoujard 1990:16), at least from a phonological point of view (cf. 2.3.1 for discussion).

1.2 The tools of prosodic representation

In order to be able to analyze and characterize the distribution of vocoids and the phonological processes affecting them, it is necessary to refer not only to segmental, but also to suprasegmental structure, i.e. to phonological constituents larger than the segment.

The elements that are introduced in this section and the next are those adopted in standard Derivational Theory (DT); although many of them provide a basis for Optimality Theory (OT), too, some of them, and mainly those which are to do with phonological derivations, are rejected by the latter (see 1.5). They all belong to different subtheories (i.e. autosegmental phonology, feature geometry, lexical phonology and, in 1.3 and 1.4 respectively, syllable theory and the metrical grid), which will be briefly introduced in separate subsections. However, such subtheories are nothing but the modules into which the (DT) phonological component is articulated, and interact freely within it.

---

[8] The issue of allophonic variation, especially as regards the quality of vowels, is not relevant to the present discussion, and will not be pursued any further. For an inventory of the allophones of Modern Standard Arabic vowels as pronounced by Iraqi speakers, see Al-Ani (1970:23). For a far more accurate account of vowel phonemes and allophones according to their environment with respect to stress placement and syllable structure, see Harrell (1957:45-62). Also see, among others, Cantineau (1960:passim) on Classical Arabic, Abumdas (1985:41-9) on Libyan Arabic.

1.2.1 Autosegmental representations

As it seems to be the best supported, most flexible and most widespread apparatus in contemporary phonological theories, this thesis adopts an autosegmental model of representation (Goldsmith 1976, McCarthy 1979a, 1981, Clements and Keyser 1983), i.e. a model in which the various segmental and suprasegmental features are placed on different tiers. Each tier is autonomous from the others, but the whole structure is kept together by association lines linking particular units on one tier to particular units on another tier. The linking procedure is regulated by the *well-formedness conditions* (WFCs) given below.

(3)

(a) Lines of association must not cross.

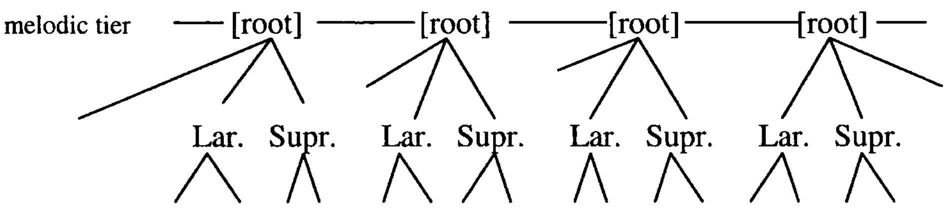
(b) All appropriate segmental slots must be linked with an autosegment.

(c) No autosegment may be left unassociated with an available and appropriate slot.[9]

(from Goldsmith 1976)[10]

As was mentioned in 1.1, at the segmental level a feature geometry model will be assumed. This is essentially a development of autosegmental phonology, in that distinctive features are not arranged into unordered matrices, but into hierarchically structured tree diagrams which can be made to hang off the melodic (i.e. the segmental) tier (cf. (4)).

(4)



[9] Conditions 2 and 3 do not necessarily imply a ban on one-to-many associations; in autosegmental phonology, long segments are represented as single autosegments associated with two (or three) timing slots (see ch. 3), while complex segments like affricates or pre-nasalised consonants are formed by two autosegments linked to the same timing slot.

[10] Goldsmith's (1976) *Universal Association Convention* ('Match the tones and tone-bearing units one to one, left to right') referred only to the association between the tonal and melodic tiers in specific tone languages. It was McCarthy (1979a, 1981), and especially Clements and Keyser (1983), who extended the use of autosegments to other phonological elements.

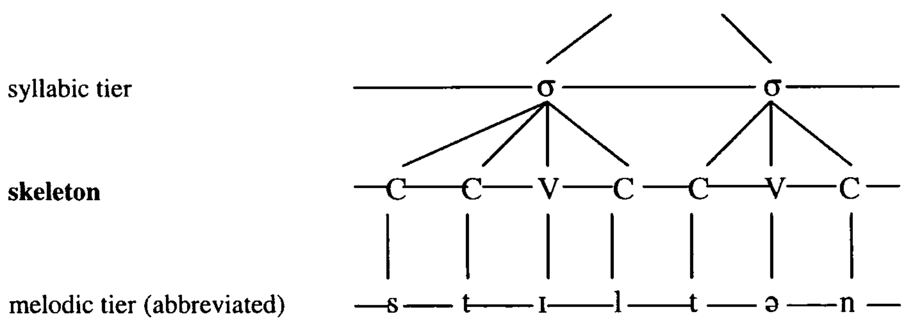
However, in this thesis feature trees will mostly be abbreviated into cover symbols (generally from the IPA) on the melodic tier, since our specific interest centres on quantitative, rather than qualitative, alternations, and thus more on suprasegmental than on segmental structure.

Apart from the melodic tier with its abbreviated feature trees, the model includes different prosodic tiers arranged into a hierarchy from the smallest to the largest (cf. 1.2.2.1 below) and a *skeletal tier* or *skeleton*, formed by a sequence of timing units to which both the elements on the melodic tier and the lowest elements in the prosodic hierarchy (i.e. the syllables)[11] are attached.

1.2.1.1 The skeleton

The skeleton is the pivotal tier: all other tiers - namely the melodic tier (or tiers, see 1.2.5) with its feature geometry trees and the syllabic tier with the prosodic tree on top - are linked to it, and through it (see (5) below). It is composed of timing slots, i.e. of a series of positions which correspond to units of time.

(5) Autosegmental representation of the word *stilton* (simplified):



The skeleton has been represented in essentially two different ways: as a CV-tier and as an X-slot tier.[12] The CV-tier model was developed by McCarthy (1979a, 1981) and Halle and Vergnaud (1980), and represents the skeleton as a sequence of V (vocalic, or syllabic) and C (consonantal, or non-syllabic) positions, to which the segments on

[11] As we will see in section 1.3, what is attached directly to the skeleton is actually the subsyllabic units rather than the syllabic nodes themselves.

the melodic tier are associated according to their feature content. In the diagram in (5), for instance, a consonant like [s] is linked to a C-slot, a vowel like [ɪ] to a V-slot. Whether glides should be assigned C- or V-slots on the CV-tier is clearly debatable. If they are linked to V-slots, the skeleton must include the feature [consonantal], while if they are linked to C-slots the skeleton must encode a certain amount of syllabicity, and perhaps include the feature [syllabic], as in McCarthy (1981). In his approach, a vowel would be represented as in (6a), a glide as in (6b).

```
(6) (a)    V              (b)     C
           |                      |
        [-cons]                [-cons]
```

The X-slot skeleton, first put forward by Kaye and Lowenstamm (1984) and Levin (1985), is instead formed by a series of pure timing slots completely lacking in feature content, and represented as 'X'. This latter model has the advantage of accounting for the cases in which the same timing slot may be associated either with a vowel or with a consonant. In Tiberian Hebrew, for instance, the final slot of the definite article is occupied by the lengthening of the previous vowel when the following stem begins with a guttural consonant (e.g. *haa-'iir* 'the city'), by a copy of the stem-initial consonant when this is not guttural (e.g. *ham-melek* 'the king') (Kenstowicz 1994:426). More generally, Compensatory Lengthening, i.e. the phenomenon by which a vowel lengthens in order to occupy a slot that was freed by the deletion of a consonant (see ch. 3), is difficult to account for in terms of a CV-skeleton, since an originally consonantal slot has to associate with a vocalic autosegment.[13] Moreover, glide ~ vowel alternations (see ch. 2) and surface syllabic consonants are more complex to explain within the CV model, in that V-slots should be allowed to link to marginal positions and C-slots to link to nuclear positions in certain circumstances. Alternatively, a C-slot should be allowed to turn into a V-slot and vice-versa, which is rather absurd.

[12] As will be shown in 1.3, a third approach, moraic theory, dispenses with an independent skeletal tier altogether.

[13] A similar phenomenon is the vocalization of floating matrices after the deletion of a consonant (see yers in Slovak, in Rubach 1993a).

The X-slot model is also preferable at the theoretical level, because the CV-tier duplicates information on the syllabicity of segments which is already stored at the syllable level.

However, as will be shown in 1.2.5 below, there are languages which seem to require skeletal templates consisting of fixed CV-slot sequences, while in other languages the same segment appears to alternate for syllabicity in certain words but not in others (see Berber in 2.1).[14] More generally, in each language some segments seem to be underlyingly non-specified for syllabicity, while others can only occupy certain positions in the syllable. The question arises of whether this is better represented in terms of language-specific constraints or diacritics, or in terms of CV- vs. X-slots. In the latter case, the possibility of adopting a CVX-tier must be taken into consideration (see 2.1 for discussion).

1.2.2 Prosodic structure

In order to deal with suprasegmental structure, a Prosodic Phonology perspective will be adopted. This theory was first proposed by Selkirk (1980a,b, 1981a) and further developed especially by Nespor and Vogel (1982, 1986), who formulated the so-called 'standard theory'.

1.2.2.1 The prosodic hierarchy

Prosodic Phonology is the phonological subsystem which deals with the formation, distribution and function of prosodic constituents, i.e. of the suprasegmental chunks into which a string of speech is divided. Each prosodic constituent is the domain of application of phonological processes, as well as a domain inside which stress assignment takes place. Such constituents, as identified in 'standard theory',[15] are seven, and are arranged in the following hierarchy:

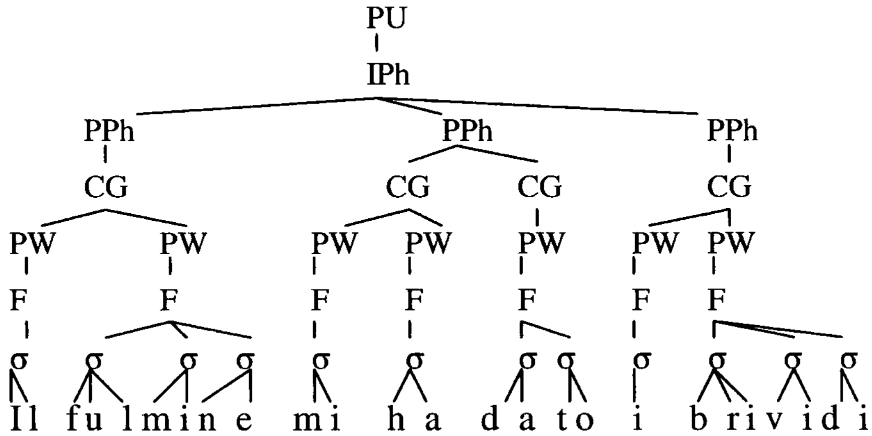
Syllable (σ) – Foot (F) – Phonological Word (PW) - Clitic Group (CG) - Phonological Phrase (PPh) - Intonational Phrase (IPh) – Phonological Utterance (PU)

[14] See Carr (1993:208) for a further argument in favour of the CV-tier, based on the fact that it allows representation of an empty C-slot to block liaison and elision before an *h-aspiré* in French.

[15] The actual composition of the prosodic hierarchy has been questioned in different ways, especially as regards the existence of certain phonological constituents. For example, the issue of whether the Clitic Group is a constituent in its own right has been raised several times (cf. James 1988:164), while Selkirk (1984a: 31) denies the existence of the Foot as a prosodic unit.

Such a hierarchy is separate from – and to some extent independent of – the morphosyntactic hierarchy, since, according to Nespor and Vogel (1986, henceforth NV), morphosyntactic units are inadequate as domains of application of prosodic rules.[16]

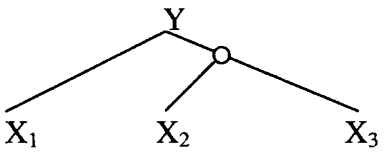
Prosodic structure can be represented by means of tree diagrams. Such graphs, as can be seen in (7), are not binary, but n-ary branching: each upper level unit X^n is composed of one or more immediately lower level units X^{n-1}, without a limit of number.



'The lightning gave me the shivers'

The authors argue in favour of n-ary branching on the grounds that binary branching – adopted by many scholars, among whom Liberman and Prince (1977) and Selkirk (1980a, 1981a) – creates unlabelled nodes which do not correspond to any constituent, as can be seen in fig. (8).

(8)



(NV:9)

[16] NV (2) point out that the two hierarchies 'differ not only in the way they divide a given string into constituents', but also 'with respect to depth'. The depth of phonological structure is finite, since the mapping rules that construct it are non-recursive, while the depth of syntactic structure is in principle non-finite, since its formation is regulated by recursive rules.

However, in more recent works n-ary branching has mostly been discarded in favour of binary branching, at least for the constituents below the PW.

Another problematic feature of NV's prosodic trees is that each lower level unit is exhaustively contained in the immediately upper level unit, so that no constituent of level X^n can attach directly onto a node of level X^{n+2} or higher. This is called the Strict Layer Hypothesis (SLH), and comprises the two requirements reported in (9).

(9) SLH:

(a) A constituent of a given level may only dominate constituents of the immediately lower level (Strict Succession).

(b) A constituent of a given level must be exhaustively contained into the constituent that directly dominates it (Proper Bracketing)

However, strict layering has been attacked on many sides, or simply ignored, so that many linguists (e.g. Prince and Smolensky 1993, Hammond 1997) allow lower level constituents to attach directly onto a non-contiguous larger constituent. Other scholars, such as Itô and Mester (1992) and Vogel (1999), have tried to save the SLH by modifying it only partially. The main problem the authors encounter is the presence of material which must be grouped into a constituent of level X even if it does not meet the minimality requirements to be labelled as a constituent of level X-1.[17] There are two possible solutions to this and similar problems, both of them somehow weakening the theory:

1. Modifying the SLH
2. Permitting the existence of subminimal constituents

Both Itô and Mester (1992) and Vogel (1999) select the first option as the less harmful to the theory: (minimal) violations of condition (9a) are permitted, while condition (9b) is kept in force.[18]

---

[17] Vogel (1999) shows for example that in Italian a monomoraic prefix cannot form a PW on its own, since it is smaller than the minimal word, which corresponds to a bimoraic F; neither can it be joined into a PW together with the following stem (see also NV:125-33). Clitics are also a problem in this sense, while monomoraic suffixes can be joined with the stem at the PW level without meeting the minimality requirements to form a F (Vogel 1999:8-9).

[18] In the case of Italian, monomoraic prefixes and clitics are allowed to attach directly at the level of the CG, monomoraic suffixes at the level of the PW (Vogel 1999:9).

A softer, and thus more realistic version of the SLH is that of *Prosodic Licensing* (Itô 1986), by which, in order to be phonetically realized, all phonological constituents must be exhaustively contained in higher, but not necessarily 'immediately higher', prosodic structure (see 1.3 and 1.4).

1.2.2.2 Prosodic rules

The processes which apply inside the prosodic domains are characterized by means of 'prosodic rules', and are purely phonological in nature. There are three types of prosodic operations, which can be represented by the following rules,[19] where A and B are segments, X, Y and Z are strings of segments and Di and Dj are prosodic categories (cf. Selkirk 1980a:111-112):

1. Domain span rules, applying within a constituent:

   $A \rightarrow B$ / [... X ___ Y ... $]_{Di}$

2. Domain juncture rules, applying at the juncture between two constituents:

   a. $A \rightarrow B$ / [... [... X ___Y $]_{Dj}$ [Z ... $]_{Dj}$ ... $]_{Di}$

   b. $A \rightarrow B$ / [... [... X $]_{Dj}$ [Y ___ Z ... $]_{Dj}$ ... $]_{Di}$

3. Domain limit rules, applying at the edge of a constituent:

   a. $A \rightarrow B$ / [... X ___Y $]_{Di}$

   b. $A \rightarrow B$ / [X ___ Y ... $]_{Di}$

Although prosodic units are not necessarily isomorphic to the corresponding morphosyntactic constituents, they are built on them by means of so-called 'mapping rules'. These latter, representing the processes by which prosodic constituents are formed, refer not only to phonological, but also to morphological, syntactic and semantic information. For example, the PPh domain is constructed on the basis of syntactic notions. NV (168) formulate its definition in terms of the following principle:

---

[19] As will be shown in more detail in section 1.3 and in chapter 5, Itô (1989) suggests that rules of this type, which she calls 'skeletal slot rules', are inadequate to characterize prosodically driven processes such as epenthesis. In her Prosodic Theory, such a process is derived directly from the syllabification principles of *directionality* and *maximality*.

> The domain of PPh consists of a CG which contains a lexical head (X) with at least one positive specification according to the categorial feature system, and all CGs on its non-recursive side up to the CG that contains another head outside of the maximal projection of X.

Once a prosodic constituent has been defined, however, access to morphosyntactic structure is precluded: prosodic rules - i.e. 'phonological rules proper' – are blind to non-phonological information and can only refer to it indirectly, through the prosodic domain inside which they apply (see also Selkirk 1981b, 1984a). In fact, NV (34-58) point out that phonological factors such as syllable count, length of a constituent and intonation contour are more relevant to the application of prosodic rules than morphological factors such as affixation or syntactic notions like those of trace and c-command.

These considerations concern one of the main issues that contemporary phonological theories are addressing, namely that of the interface between phonology and grammar (in the sense of 'morphosyntax'): is phonology completely grammar-driven, or does it, at least in some cases, act as an independent component with its own hierarchy of constituents, different and independent of the morphosyntactic hierarchy? The evidence NV produce in favour of this latter hypothesis appears to be compelling, while other scholars, such as Rischel (1987) and Fudge (1999, see 1.4) even go so far as to postulate the existence of more than one prosodic hierarchy.

Semantics also comes into play, especially when dealing with the largest postlexical constituents (IPh and PU, and perhaps PPh) and with connected speech phenomena. In NV's model, the structuring and restructuring of IPh and PU is directly influenced by semantic contents. Functional-semantic factors such as sentence perspective, especially in 'freer order' languages like Italian and Arabic, can in fact have a strong impact on the organization of the prosodic subsystem. Intonational patterns, which are also fundamental in determining the shape and extension of the larger constituents and the application of prosodic rules, are often constructed on the basis of semantic information. In its turn, the semantic component sometimes seems to require direct access to phonological material. NV (249-271) argue, for example, that disambiguation works on the basis of phonological structure rather than of syntactic structure. There must therefore be a direct phonology/ semantics interface, without syntax coming in as a mediator.[20] This is in contrast with standard SPE (Chomsky and Halle 1968) theory, which sees both phonology and semantics as unconnected outputs of the syntactic component.

1.2.3 Interaction with morphology and syntax: the lexicon

1.2.3.1 Lexical phonology

As was mentioned in 1.2.2.2, Prosodic Phonology is interested mainly in purely phonological processes, unaffected by morphosyntactic information. However, in order to characterize the distribution of segments, it is necessary to look at the interaction between phonological and morphological structure, too. Within Derivational Theory, Lexical Phonology, first introduced by Kiparsky (1982) and Mohanan (1982), provides ample means to do exactly that. It is a theoretical framework that analyzes phonological processes which are sensitive to morphological structure as taking place within the lexicon, along with morphological processes of affixation with which they are interleaved, and therefore preceding all other, *postlexical,* processes. Lexical operations exhibit certain properties that differentiate them from postlexical operations. These are:

a. While postlexical processes are blind to grammatical structure, lexical processes are sensitive to it.

b. While postlexical processes are 'natural' in that they promote ease of articulation, lexical processes are generally phonetically unmotivated.

c. While postlexical processes are exceptionless, lexical processes often have exceptions.

d. While postlexical processes are regular and productive, lexical processes are idiosyncratic and non-productive.

e. While postlexical processes may introduce allophones that are not present at the underlying level, lexical processes are *structure preserving*, i.e. they produce alternations between underlying phonemes.

[20] In Selkirk's (1984a) revised prosodic theory, on the other hand, the relation between sound and meaning is construed as mediated by syntax. However, not only surface, but also deep syntactic information is relevant.

f. While postlexical processes are often optional, lexical processes are always obligatory.

g. While postlexical processes may be gradual, lexical processes always apply in an on-off fashion.

Lexical (phonological and morphological) processes are further subdivided into different strata according to the principle of *level ordering*, so that all processes belonging to level 1 apply before all processes belonging to level 2, and so on. Moreover, once all the morphological and phonological operations at a certain level are completed, the mechanism becomes blind to the internal structure of such a level. This is accomplished by means of a principle of *bracket erasure*, by which internal brackets are erased at the end of each level. Kiparsky (1982:152) further points out that certain lexical processes only apply in a *derived environment*, i.e. in an environment that is the result of a morphological or phonological process.

1.2.3.2 Lexical prosodic categories

One of NV's basic assumptions is that lexical phonology precedes prosodic phonology, which leads to the prediction that no lexical rules will be able to refer to prosodic units such as the σ, F and PW. Such a notion has been questioned by scholars such as Booij (1988), Inkelas (1989) and Booij and Lieber (1992), who show that not only lexical phonology, but even some lexical morphology rules must have access to the prosodic constituents below the PW. For instance, the English comparative suffix *–er* can only be attached to monosyllabic adjectives plus a certain group of disyllabic adjectives, which implies that this kind of morphological formation must have access to the prosodic notion of σ. We can thus conclude that, while the constituents above the word are entirely postlexical, σ, F and PW are already formed inside the lexicon (cf. Booij 1988).

This does not, however, imply that late postlexical rules may not have these 'lexical' constituents as their domain of application. As Wiltshire (1998:425) puts it, 'the domain of a rule or constraint and its level of analysis need not be conflated. The edges of smaller domains do not necessarily take precedence over the edges of larger domains.'

1.2.3.3 Phrasal allomorphy

Another problem with NV's approach is the presence of 'direct syntax' rules, i.e. of phonological rules which refer directly to syntactic structure in their formulation. Such rules have been introduced in order to account for phonological processes which cannot be analyzed in terms of prosodic domains, since they apply only to certain specific lexical items or syntactic categories (cf. NV:28-33; Nespor 1993:217). For example, the French singular feminine possessives *ma* 'my', *ta* 'your' and *sa* 'his/her' become *mon*, *ton* and *son* respectively when followed by a vowel-initial word inside the same noun phrase. The phenomenon is clearly phonologically motivated, but it is impossible to formulate it as a prosodic rule since it applies only to a restricted set of words (see Nespor 1993:218-9). In this case, NV would allow for a 'direct syntax rule' to apply. The presence of such rules, however, obviously weakens the theory: if the domain of application of a rule cannot be identified with a specific prosodic constituent, one can always resort to the direct influence of syntax. In order to solve the problem, Hayes (1990) suggests that in these cases one may be in the presence of a kind of 'phrasal allomorphy', i.e. of phrasal phonology precompiled in the lexicon. This means that phonological information can already be contained in the subcategorization frames of single items inside the lexicon; the more specific environments will be defined first, while the Elsewhere Condition will provide the form to be inserted everywhere else. In the case of the French possessives, the frames will be defined as in (10):

(10) Feminine singular possessive in French:
mon, ton, son / ____ V
ma, ta, sa / elsewhere

(Nespor 1993:220)

1.2.4 Fast speech rules

Apart from the *prosodic*, or *connected speech* processes, which can be dealt with in terms of prosodic domains, another, perhaps less interesting, class of postlexical rules has often been mentioned in the literature (see for example Kaisse 1985, NV). These are called *fast speech rules* because their implementation, which is always optional, is promoted by an increase in speech tempo. Just like prosodic rules, they are purely phonological in nature, but, contrary to the latter, they are not restricted to apply within a definite prosodic domain. Their application is triggered mainly by ease of articulation, and depends only on factors such as speech rate and sentence perspective. Besides, contrary to prosodic processes, fast speech rules may lead to phonotactically disallowed sequences, as in the pronunciation of the word *potato* as [p^{h}t^{h}ɛ^{j}ɾə] following the application of Vowel Deletion in American English (NV:23-4). According to Kaisse (1985:19), who however includes among them connected speech phenomena such as American English flapping, these processes belong to a 'second postlexical stratum'.

1.2.5 Arabic and non-concatenative morphology

Italian, like the other Indo-European languages, exhibits a *concatenative* morpho-phonological system, i.e. its morphemes are attached to one another in linear order, as in *in-cred-i-bile* 'incredible' or *ri-struttur-a-va* '(s/he) was restructuring'. Semitic Arabic, on the other hand, is a *non-concatenative* (or *root-and-pattern*) language, in which morphemes can be discontinuous. In particular, consonantal roots appear to be interspersed among vowel patterns that vary from one derivative to another. The Classical Arabic words in (11), for instance, clearly share one and the same root *ktb*, which gives them the common meaning of 'writing'; however, they all have different vocalic patterns (sometimes accompanied by affixal consonants and/or reduplication), which function as discontinuous affixes.

(11) ***kataba*** 'he wrote'

kattaba 'he caused to write'

kaataba 'he corresponded'

takaatabuu 'they kept up a correspondence'

ktataba 'he wrote, copied'

kitaabun 'book (nom.)'

kuttaabun 'Koran school (nom.)'

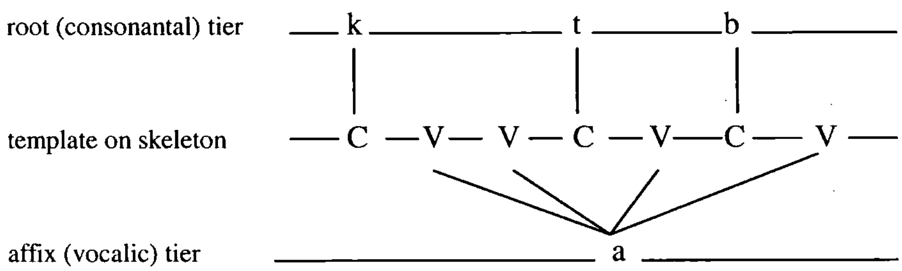
kitaabatun 'act of writing (nom.)'

maktabun 'office (nom.)'

(McCarthy 1981:374)

McCarthy (1979a, 1981) first had the intuition of locating root and affixal segments on two different tiers, which are later conflated to constitute words. In his model, the pattern is instead given by prosodic templates on the skeleton. Such templates are sequences of C and V positions: the root consonants are mapped onto C-slots and the affixal vowels onto V-slots.[21] For instance, the word *kaataba* 'he corresponded' is represented as in (12).

(12)



In many cases, a certain template corresponds to a specific grammatical category, such as a verb tense or aspect.

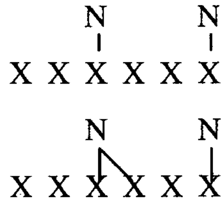
McCarthy makes use of a CV skeleton rather than an X-slot skeleton because consonants and vowels appear to be placed on different tiers, and must be able to see if a slot is C or V in order to be placed in the correct order. However, even though affixal segments are typically vowels and root segments are typically consonants, vocoids may actually be placed on either tier, according to whether they are part of a root or of an affix. Even more interestingly, as will be shown in section 2.3, later phonological operations may create vowel ~ glide alternations on both tiers. An X-slot skeleton would clearly be more suitable to portray such a situation. In X-slot theory,

[21] According to McCarthy (1981), the templates are composed of the features [segmental] and [syllabic].

the distinction between ‘consonantal’ and ‘vocalic’ slots can actually be maintained by pre-linking the latter to nuclear positions underlyingly, as in (13).

(13) (a) form II verbs template: CVCCVC=

(b)form III verbs template: CVVCVC =



(from Watson 2000:2)

However, both the CV-tier and the pre-linked X-slot models encounter the same difficulty with respect to vocoids: nothing would prevent a high vocoid in the root from attaching to a V- (or pre-linked) slot or one in the affix from attaching to a C-slot, unless an awkward distinction between ‘consonantal’ and ‘vocalic’ vocoids is introduced.

Since McCarthy’s tiers are morphologically defined, a possible solution seems to be that of distinguishing the slots of the templates in morphological, but not in phonological, terms. The templates would thus be composed of X positions marked as R (root) or A (affix), rather than of C- and V-slots. An obvious drawback of this approach is the question of how to account for the fact that the roots are made up by segments that can (and generally do) occupy marginal positions within the syllable, while the affixes are made up by segments that generally occupy nuclear positions.

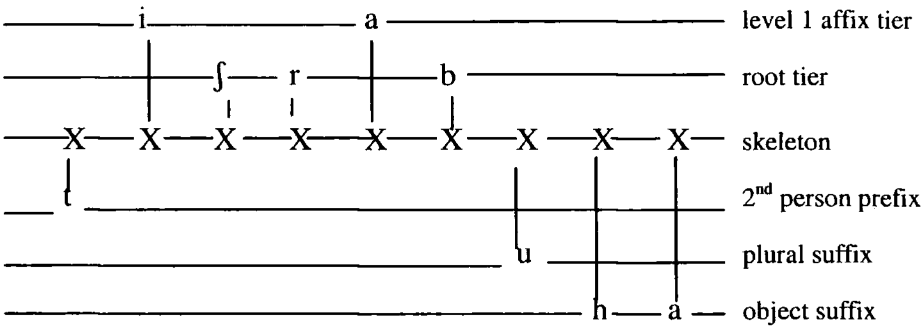
The vocalic affixes that have been mentioned so far are what may be termed *level 1 affixes*, and form the stem of Arabic words together with the consonantal root. There are, however, other types of affixes, such as prefixes and suffixes, that are concatenative and which, in a lexical morpho-phonological framework, are assigned to different levels. These, combined with the stem, form the Arabic grammatical word. For instance, inflectional endings (i.e. subject agreement suffixes) belong to level 2, object pronominal clitics and possessives to level 3.[22] More precisely, Kiparsky (2000), following Kenstowicz (1983), points out that the subject agreement,

[22] For arguments in favour of such a categorization, based on the interaction between stress assignment, syncope and epenthesis in Palestinian Arabic, see Brame (1974), Kenstowicz (1981), Kiparsky (2000) and especially Kenstowicz (1986). Also see sections 4.3 and 5.3 of this thesis.

feminine, plural, dual and 'singulative'[23] suffixes are attached at the stem level, while the object and possessive suffixes, the postposition *-l* 'to' and the negative marker *-sh* are attached at the word level.[24]

Each of these morphemes is allocated a separate tier (see 14), because the tiers are morphologically, and not phonologically, defined (McCarthy 1981:387).

(14) AA *tishrabuha* 'you (pl.) drink it (f.)'



This is further confirmation that, as was put forward above, there are really no 'consonantal' and 'vocalic' tiers, since not all consonants belong to the same tier and some tiers host both vowels and consonants. For example, the consonant [h] of the level 3 suffix *–ha* 'it (f.)' is placed on a different tier from the root tier, together with the following vowel.

However, the secondary importance of vowels in Arabic with respect to consonants is clearly a reflection of their morphological affiliation. Since they are typically affixal components, vowels do not seem to enjoy the same status as consonants in the minds of native speakers; this is probably both the cause and the result of the fact that the former are not represented in the spelling. In fact, Arabic vocoids tend to alternate and undergo changes more readily than contoids, which generally form the root of words and carry more specific meanings. As will be shown in the following chapters, in some cases phonological processes apply to vocoids in order to avoid the loss of a consonantal sound.

---

[23] Singulative suffixes make singular nouns out of mass nouns, e.g. *ba'ar* 'cattle' > *ba'ara* 'a cow'.

[24] According to Abu-Mansour (1992:57), however, in Makkan Arabic suffixes are assigned to three distinct levels: agreement suffixes belong to level 1, prepositional clitics to level 2, object pronouns to level 3 (see 3.3 for evidence in favour of this hypothesis).

1.3 Syllable structure

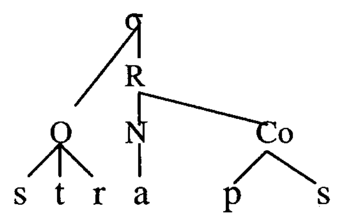
As was mentioned in section 1.1, syllabicity is what distinguishes vowels from glides, and most of the definitions and categorizations related to vocoids are based on this notion. For this reason many phonological processes affecting vocoids are triggered or blocked by syllable structure constraints. It is therefore essential to take a closer look at syllable structure and the mechanisms governing it.

This section will be looking first at the internal make-up of syllables and at syllable structure constraints, then at the process of syllabification, by which segments are parsed into syllables. The last two subsections are dedicated to syllable structure in Italian and Arabic respectively.

1.3.1 Subsyllabic units

The syllable is the smallest unit in NV's prosodic hierarchy, i.e. the minimal domain of application of phonological rules. However, it can be broken up into smaller units, the onset (O) and the rhyme (R), which in turn consists of a nucleus (N) and a coda (Co), as in the following example:

(15)



Such constituents can be motivated on the basis of the role they play in phonotactic constraints and stress assignment. However, they cannot be considered as part of the prosodic hierarchy for a number of reasons (NV:12-13). First, these units are of different types, even if they are all grouped under the same syllable node. Secondly, while the other prosodic domains can contain one or more subordinate units, the syllable cannot contain more than one onset, or more than one coda. A third piece of evidence for the exclusion of syllable internal constituents from the prosodic hierarchy is so-called 'ambisyllabicity', whose existence has largely been proved in metrical theory. When a segment is ambisyllabic, it belongs to the rhyme of one syllable and to the onset of another syllable at the same time. This is in open contrast with the Proper Bracketing condition of the SLH (see 1.2), thus enforcing the notion that sub-syllabic units are not prosodic constituents in their own right. Last but not least, such units also differ from the other prosodic constituents in that they do not serve as the domain of application of phonological operations.

The arrangement of segments within the syllable is related to their placement in the sonority hierarchy (cf. Angoujard 1990:16-24), which is essentially to do with the degree of stricture involved in producing them: low vowels, being more sonorous, are in fact the most likely elements to occupy the nucleus, voiceless stops the least likely (see 1.1). Moreover, the segments in the onset and coda also adhere to the sonority scale, with the most sonorous elements being closer to the nucleus (cf. Selkirk's (1984a:16) *Sonority Sequencing Principle*).[25]

In all syllables, the nucleus position is obligatorily filled, while onset and coda can be empty. However, in many languages there is a ban on onsetless syllables (V or VC), especially word-internally, and even in the languages where onsetless syllables are allowed, the formation of onsets is highly favoured. That is, an intervocalic consonant will always be syllabified together with the following vowel, and never with the preceding one. In DT, this property is called the *Onset Principle*, and is generally formulated as a constraint, or a preference, such as the one in (16).

(16) *Onset Principle*

Filled onsets are preferred.

or

$* _{\sigma}[V$ (Avoid syllable-initial vowels, or better, nuclear elements).

Such a principle is universal but parametric, in that it can be either 'relative' or 'absolute', according to whether a language allows for onsetless syllables when no onset candidate is available (Itô 1989:223).[26]

[25] According to Dell and Elmedlaoui (1985) this is not always true, and in some languages, such as Imdlawn Tashlhiyt Berber, other syllable structure constraints take priority over Sonority Sequencing.
[26] In reality, this constraint is gradient, i.e. it has more than two degrees of implementation. This seems to point towards an OT analysis of the issue, with the constraint being violable (cf. 1.5).

Moreover, a tendency not only to fill, but also to maximize onsets is widespread, especially among Indo-European languages (Kenstowicz 1994:257). This can be expressed in terms of the so-called *Onset Maximization Principle* (17):

(17) Syllabify a consonant into an onset whenever this does not violate sonority sequencing constraints.

A connected tendency is the preference for codaless (i.e. open) syllables, or the requirement that a coda consonant be homorganic to the following onset consonant. This is what Itô (1989) calls the *Coda Filter*, by which only doubly place-linked consonants are allowed in the syllable coda (18).

(18) *Coda Filter*

$$* \; C]_{\sigma}$$

|

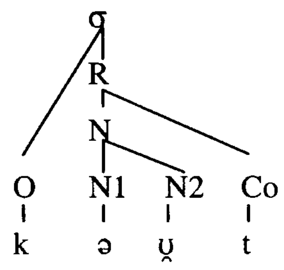
[PLACE]

The Coda Filter is common to many languages, among which Japanese, Ponapean, Diola Fogny, Lardil. Moreover, modified coda filters are found in Finnish, Italian and even English (Itô 1989:225).
Cross-linguistically, complex onsets and codas are also disfavoured. It should be noted that the preference for simple onsets can actually counteract Onset Maximization when a sonority increasing 2-consonant cluster is to be syllabified: in some languages, it will be parsed as a complex onset, in others as a coda + simple onset. Itô (1989:219-220) accounts for such a difference in terms of a directionality of syllabification parameter, as opposed to an account based on the extrinsic ordering of an 'onset rule' maximizing onsets and a 'coda rule' maximizing codas (so long as the requirement for filled onsets is fulfilled). When the direction of syllabification is right-to-left, a complex onset is formed, while when the direction of syllabification is left-to-right the first consonant is syllabified into the coda of the preceding syllable. Although Onset Maximization is parametric, 'onset satisfaction', i.e. the Onset Principle, is universal. Besides, conditions such as the Onset Principle and the Coda

Filter do not fall out of the directionality parameter, and must be stipulated separately (Itô 1989:222).[27]

Since syllable structure is allocated a different tier to segmental structure, a given segment can in principle occupy any position inside a syllable. In reality, in the vast majority of languages only the most sonorous elements can be placed in the nucleus. The actual presence of a segment in such a position depends on language-specific phonotactic constraints, and on the sonority of the adjacent segments. By virtue of their high degree of sonority, vocoids are obviously most likely to be assigned to the nucleus. In many languages, in fact, they are the only possible nuclear elements. Less sonorous vocoids such as /I/ and /U/ can of course occupy non-nuclear positions when they are adjacent to more sonorous segments. Such non-nuclear vocoids are traditionally called 'glides' (cf. 1.1). However, off-glides, on the grounds of their vowel-like quality, have often been analyzed as placed in the syllable nucleus together with the preceding vowel (O'Connor 1971, Lass 1976, Gimson 1989, Mioni 1993, Harris 1994), rather than in the coda (cf. 2.1).[28] A monosyllable like *coat* [kəʊ̯t] can thus be represented as follows:

(19)



1.3.2 The mora

Syllable types can also be classified according to their weight, often expressed in *moras* (see Hyman 1985, Mc Carthy and Prince 1986). Originally, a mora (μ) is in fact a unit of quantity which is used to formalize the notion of weight, although so-called Moraic Theory views moras as veritable prosodic positions. For most languages, a light syllable is monomoraic, a heavy syllable bimoraic and a superheavy has three or

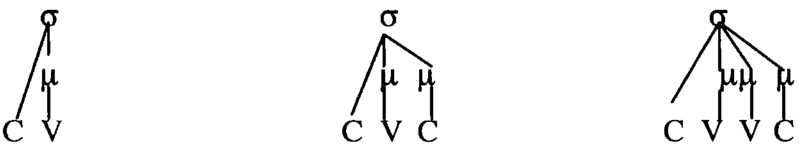
[27] In section 1.5 it will be shown how all of these preferences, being apparently present in all languages but implemented with different degrees of strength, can be straightforwardly represented in terms of OT constraints (see Prince and Smolensky 1993) rather than in terms of rules and bipolar parameters.

[28] In chapter 2 it will be argued that, in languages like Italian or Spanish, even on-glides may belong to the rhyme (see Harris 1983, Harris and Kaisse 1999).

more moras. A short vowel is assigned one mora, a long vowel two moras; each coda consonant is also assigned one mora, while onsets are weightless. However, while vowels and geminate consonants exhibit a mora underlyingly, simple consonants do not: they are assigned a mora only when they appear in a coda, by *Weight-by-Position* (Watson 2000:3). Hence, a light syllable will be of the shape V or CV (or CCV, etc.), a heavy syllable of the shape (C)VV, (C)VG (where G = glide) or (C)VC, a superheavy syllable of the shape (C)VVC, (C)VCC, (C)VCCC etc.

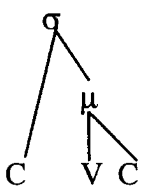
Syllable weight can be represented on a moraic tier, as follows:[29]

(20)



In some languages, however, CVC syllables are considered light, i.e. monomoraic. They can thus be represented as follows:

(21)



Hayes (1989) recapitulates the sources of syllable weight as follows:

(a) Short vowels contribute one mora and long vowels two moras (universal)

(b) Geminate consonants contribute one mora (universal)

(c) Weight-by-Position: a 'coda' consonant is assigned a mora in the course of syllabification (parametric)

[29] In Itô's (1989) framework, onset consonants are not adjoined directly to the syllable node, but form a mora together with the nuclear vowel:

i.

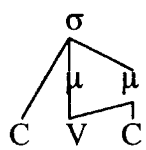
σ

μ μ

C V C

Cross-linguistically, trimoraic syllables are dispreferred, if not altogether avoided by means of devices such as insertion, deletion or shortening (McCarthy and Prince 1986, Steriade 1991). While peripheral CVVC or CVCC syllables can be considered as bimoraic syllables with a final extrasyllabic consonant (see 1.3.3.2 below), non-peripheral CVVC have been represented as bimoraic by means of a device called 'Adjunction to Mora' (see Broselow 1992, Broselow et al. 1995, Watson 1999a, 1999c and 2002), by which a CVVC syllable is represented as follows:

(22)



The fact that monomoraic syllables are also often avoided – even if never completely banned – has led to the formulation of the Bimoraicity Constraint in (23).

(23) *Bimoraicity Constraint:*

All syllables are maximally and optimally bimoraic.

(Broselow 1992).

Such a constraint, however, is not universal, since there are languages, such as Hindi, which have a clear three-way weight distinction (Broselow et al. 1995).[30]

As will be shown in section 1.4, syllable weight often determines the distribution of stress, in that heavy syllables are more readily accented than light syllables.

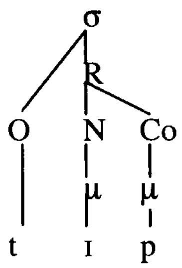
Having introduced a moraic tier, Moraic Theory abandons the skeleton, thus claiming to be representing only 'genuine prosodic positions' (Watson 2000:3). In fact, most supporters of this theory seem to assume that the mora is part of the prosodic hierarchy, sometimes making this claim explicit (see for example Bullock 1991), although such a unit does not meet NV's criteria for inclusion.

[30] Again, this points towards OT's violable constraints (see 1.5).

Moraic Theory also generally avoids the use of subsyllabic nodes such as onset and rhyme. According to Watson (2000:4), Moraic Theory is to be preferred to Onset-Rime Theory because it gives prosodic status only to those positions that bear syllabic weight and have influence on stress assignment.

However, even more substantial arguments can be used in favour of the Onset-Rime Theory. For instance, processes which involve the onset (such as Consonant Insertion) cannot be straightforwardly represented in a moraic framework, because such a position does not project a mora. In addition, Moraic Theory provides no means of differentiating between diphthongs and vowel-consonant sequences, while Onset-Rime Theory can place an off-glide in the nucleus and a post-vocalic consonant in the coda (cf. Watson 2000:5). Therefore, in this thesis representations will be assumed in which moraic and onset/rime features coexist, as in (24).

(24)



Such representations, however, will often be abbreviated according to what kind of structure is relevant to the argument.

1.3.3 Syllabification

The process of syllabification, i.e. the parsing of segments into syllables, has been represented and accounted for in different ways. In earlier works, such a process was formalized by means of rules applying either at a single point in the derivation (Kahn 1980) or cyclically (Steriade 1982). Mc Carthy (1979b), on the other hand, introduced the concept of 'continuous syllabification', by which segments are organized into syllables at all levels of phonological representation, with a reparsing of the string taking place at each step in the derivation.

A related, and even more debatable, issue is the question of where initial syllabification occurs. Three main views can be identified within the debate:

(a) (Complete or partial)[31] underlying syllabification, by which each word in the lexicon is already associated with syllable structure.

(b) Initial syllabification in the lexicon.[32]

(c) Surface, postlexical syllabification.

Since each of the three suggestions seems to work best in certain languages and not in others (cf. 1.3.4 and 1.3.5), the location of initial syllabification can be parameterized. A skeletal-rule-based approach to syllabification would traditionally portray the process as follows:

(a) Nucleus formation: V → [V]

(b) Onset formation: C[V] → [CV]

(c) Coda formation: V]C → VC]

However, a much clearer and more explanatory representation can be achieved when syllable structure itself is included in the diagram.

Itô (1986, 1989) argues for *templatic syllabification*,[33] by which a string is parsed directionally into syllables whose shape is characterized by a template defined both on the basis of universal principles of syllabification and of language-specific requirements. In her model, according to a parameter that is set for each language, a word is scanned either from left to right or from right to left, and first assigned moraic structure, then syllabic structure. In Arabic, for instance, a mora is first assigned to each CV pair, then to each of the consonants which are left disjoined. Finally, Arabic having maximally bimoraic syllables, the moras are progressively parsed into syllables two by two wherever possible (i.e. avoiding the presence of two nuclear elements in a syllable and the violation of sonority sequencing) and one by one in the other cases. It must be pointed out, however, that exhaustive assignment of moraic structure previous to syllabic structure is only functional when dealing with languages like

[31] For instance, Selkirk (1981b) argues for complete underlying syllabification, while Archangeli's (1984) model assumes that syllabification is only partially underlying, with syllable heads being present in the UR.

[32] According to NV (64), initial syllabification can be viewed as a universal PW-span rule (cf. also Kahn 1980:51).

[33] As will be shown in chapter 5, Itô (1989) draws evidence in favour of the templatic approach to syllabification from epenthesis, pointing out the difficulties that a rule-based approach assuming that syllable structure is a projection from the segmental level encounters.

Arabic, which only exhibit simple onsets (see 1.3.5). Languages with complex onsets, on the other hand, require syllable structure to be assigned prior to the assignment of any moras to pre-consonantal consonants, since these may be parsed either into the preceding coda or into the following onset, according to their sonority and that of the adjacent sounds. In Watson's (1999a, b, 2000) model, in fact, a mora is first assigned only to vowels. Then syllable structure is assigned, incorporating first the syllable nucleus, then the onset and finally the coda consonants, which are thus assigned a mora by position. This model, however, is still not clear in its treatment of vocoids that may alternate for syllabicity: are they, like other vocoids, assigned a mora in the first place, which is deleted when syllable structure is added if they occupy an onset position? Or are they only assigned weight later in the derivation, by position, after syllable structure has been assigned?

The model can account for the distribution of alternating vocoids (and any alternating contoids, too) if, every time a syllable is projected onto the existing skeletal structure, its subsyllabic nodes are also projected, with their relative moraic structure, regardless of whether they can be mapped onto melodic slots, or roots.[34] The subsyllabic nodes are linked to the melodic slots wherever possible, according to the parameters and phonotactic constraints set by the language. For instance, in an English word like *lachrymose* [lækrɪməʊ̯s], with syllabification applying from right to left, the final /s/ is first mapped onto the coda of the rightmost syllable. Since the /U/ cannot be assigned to such a coda, it is mapped onto the preceding nucleus, and so is the /ə/; then the /m/, which cannot be part of the same nucleus, is linked to the syllable onset. For sonority reasons, the preceding /I/ cannot be mapped onto the same onset, nor can it occupy the coda of the penult syllable; it is thus syllabified into the nucleus of the latter, while both /r/ and /k/ fit into its onset. And so on.

If the subsyllabic nodes which are left empty are not obligatory, they are deleted; if they are obligatory, readjustment processes such as syncope and epenthesis take place. However, templatic, directional syllabification does not always obtain the correct surface forms when these processes are involved (see ch. 4 and 5 for discussion). Moreover, as will be shown in 1.3.4, this approach cannot straightforwardly account for words which display hiatuses involving high vocoids, such as Italian *fantasia*

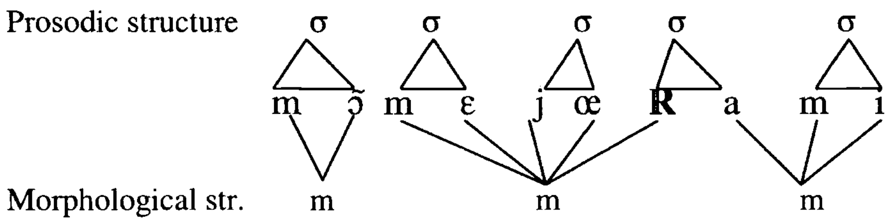
[34] For a similar model, cf. Noske (1993), who, however, adopts the CV-tier rather than moraic structure.

'fantasy, imagination'. This is because they exclusively assume surface, postlexical syllabification.

1.3.3.1 Resyllabification

Many of the languages that have underlying or lexical syllabification also exhibit a second stage of postlexical *resyllabification.* Although resyllabification applies across-the-board – i.e. the whole template is mapped again onto the whole string – it becomes evident when a segment is syllabified across a word boundary in order to fill in an empty onset. This results in a misalignment of the prosodic and morphological structure, as in the French phrase *mon meilleur ami* [mɔ̃mɛjœRami] 'my best friend (m.)' in (25).

(25)



Resyllabification is particularly common in Romance languages such as Spanish, French and Italian (see 1.3.4). NV (69-70), by giving an account of nasal + glide sequences and Nasal Assimilation in Spanish, argue that two separate rules of syllabification and resyllabification must be postulated, because there are processes that make reference to PW-level syllable structure, before resyllabification takes place. While lexical syllabification always applies within the PW, the domain of application of resyllabification appears to vary cross-linguistically. By looking at the phenomenon of Liaison, NV (72) establish that the domain for resyllabification in French is the PPh, while in Spanish such a domain is larger, probably the IPh or the PU.

1.3.3.2 An alternative: the CVX- tier

A way to obviate the need for underlying syllabification is the introduction of a CVX-tier, i.e. of a skeleton formed by both CV- and X-slots. On such a tier, C-slots are linked to segments that can only be syllabified in marginal (i.e. onset and coda) positions, V-slots to segments that can only be syllabified as syllable peaks, X-slots to segments that may alternate in their syllabicity (cf. 2.1 for exemplification and further explanation). However, this model would still assume a classification of languages into three groups: those in which syllabification takes place only within the lexicon, after word-formation, those which only have postlexical syllabification and those in which both lexical syllabification and postlexical resyllabification take place.

1.3.3.3 Extrasyllabicity

All segments must be licensed, i.e. parsed into syllables. Segments that are left unparsed are subject to *Stray Erasure* (Steriade 1982), i.e. they are not phonetically realized, unless some process such as epenthesis intervenes to rescue them. This is a corollary of the more general *Prosodic Licensing* (Itô 1986), by which all phonological units must be parsed into larger prosodic constituents.

Peripheral segments may be left unparsed because they are subject to *extrasyllabicity*, i.e. made invisible to syllabification (Clements and Keyser 1983). Extrasyllabicity is a particular case of *extrametricality*, first introduced by Liberman (1975), by which peripheral prosodic material is skipped by the metrical parse. It has proved to be useful in accounting for a large number of apparent idiosyncrasies exhibited by languages. A well-known example is French liaison, by which a word like *mes* 'my (pl.)', which is pronounced as [me] in isolation or in a pre-consonantal context, becomes [mez] when it is followed by a vowel-initial word (e.g. *me[z]amis* 'my friends') . If we analyze the final /z/ as extrasyllabic, we can say that the segment, which fails to be syllabified into the coda of the syllable [me], can be (re)syllabified into the empty onset of a vowel-initial syllable at a later stage; if there is no such syllable, the unsyllabified consonant is deleted by means of Stray Erasure.

In other languages, however, an extrasyllabic segment may also be resyllabified into the preceding, rather than the following, syllable (cf. Watson 2000:6). The segment thus escapes Stray Erasure, but, because of its late syllabification, misses the chance of being assigned a mora (see 1.3.5).[35]

[35] Selkirk (1981b), on the other hand, puts forward a model that does away with extrasyllabicity by allowing the extra consonants to be part of degenerate syllables lacking a nuclear element. The empty nuclei are later filled in with epenthetic vowels (cf. ch. 5).

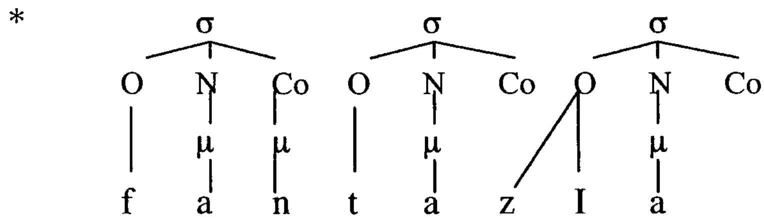
1.3.4 Syllable structure in Italian

In Italian the core syllable types are V, VC, CV, CVV, CVC, that is syllables are maximally bimoraic and minimally monomoraic. Only vowels can constitute the peak of a syllable. Onsets are not obligatory, but they can be complex. The only permitted syllable-initial consonant clusters, however, are of the form *stop* + *liquid* or *glide*. These can be preceded by a domain-initial /s/, which has been seen as extrasyllabic (see Nespor 1993). Examples of complex onsets are [**gri**:do] 'scream', [**pja**:no] 'plan', [**stra**:da] 'street', [**skwɔ**:la] 'school'. Complex codas, on the other hand, can only be found in domain-final loan words such as *trust* (in the sense of 'corporative body'), where in fact the final consonant is probably extrasyllabic. Moreover, except for loan-words such as *jet*, only the first segment of a geminate, sonorant consonants and [s] can occupy the coda of a syllable. Onset Maximization is thus widely implemented.

If we assume a templatic point of view (Itô 1986, 1989, Noske 1993), the direction of syllabification in Italian must therefore be right to left, whereby as many consonants as possible are syllabified in the onset, in accordance with Sonority Sequencing. A string like [padro:ne] 'owner' will thus be syllabified as [pa.dro:.ne] rather than *[pad.ro:.ne], while a string like [kantso:ne] 'song' will be syllabified as [kan.tso:.ne]. However, templatic approaches assuming surface syllabification encounter obstacles when dealing with Italian words which display hiatuses involving high vocoids, such as *fantasia* [fan.ta.'zi.a] 'fantasy, imagination' or *suo* ['su.o] 'her/his (m. sg.)'. In a directionality approach, there is no way to prevent the high vocoid in these words from surfacing as a glide, unless it is already marked as syllabic (see incorrect syllabification of *fantasia* in (26)).[36]

[36] Other, less desirable solutions would be to have stress assigned prior to initial syllabification, or to have the relevant high vocoids as underlyingly long (see 3.2). Otherwise, following Watson (1999a, b, 2000), all vocoids could be assigned a mora in the first place, and then high vocoids in pre-vocalic position would lose their syllabicity (and therefore their mora) because of postlexical Glide Formation (see ch. 2), unless they are stressed. However, this approach would still require subsequent resyllabification. Another possible solution would be to mark the final vowel in these words as extraprosodic, since the relevant vocoids generally appear exclusively in penultimate position (a word

(26) direction of syllabification ⇐



Hence, these data are more readily accounted for by assuming (at least partial) underlying syllabification for Italian, or by adopting a CVX-tier.

However, resyllabification applies across word-boundaries, as in *con un amico* 'with a friend (m)', which is invariably resyllabified as [ko.nu.na.mi:.ko]. Postlexical phenomena which involve reparsing of segments into syllables, such as Glide Formation and Syntactic Doubling, are in fact common in this language. According to NV (172), the domain of resyllabification in Italian is the PPh, i.e. no segments can be resyllabified across PPh-boundaries.

1.3.5 Syllable structure in Arabic

Syllable structure and the related constraints vary considerably from one Arabic variety to another, and especially between Classical Arabic and the so-called 'dialects'. Nevertheless, some generalizations can be made.

In all dialects, the core syllable types are CV, CVV and CVC,[37] but CVVC, CVCC and, more rarely, CVVCC[38] can also be found. CVCC and CVVCC are generally limited to domain-final position, so that the last consonant can be considered

like *siano* ['si.a.no] '(they) are (pres. subj.)' exhibits a derived environment, a word like *Nuoro* ['nu.o.ro] (city) is borrowed from Sardinian).

[37] According to Watson (1999b), at least in San9ani Arabic, CVC syllables differ from CVV or CVG (where G = geminate): while the latter are always heavy, CVC syllables are treated as heavy on the lower layer of a syllable-internal grid, as light on the upper layer. That is to say, while CVV and CVG syllables are underlyingly bimoraic, the coda consonant of CVC is assigned a mora only later, by Weight-by-Position. This is because CVC syllables behave as heavy in the environment of other CVC or CV syllables, as light in the environment of final CVVC or word-internal CVV or CVG.

[38] According to Abu-Mansour (1992:68), only SA has CVVCC syllables (e.g. *saarr* 'joyful', *maarr* 'passer-by'). In the other dialects, when words of this type have not altogether disappeared, word-final geminates have been simplified (this is the case in AA) and/or the preceding vowel has been shortened.

extrasyllabic.[39] Since CVVCC syllables are very rare, it has been suggested that the maximal core syllable, at least in most dialects, is bimoraic (see McCarthy and Prince 1990, Broselow 1992, Watson 1999a). In many varieties, however, CVVC syllables can be found PW-internally, so that the last consonant cannot be seen as extrasyllabic. In order to maintain the bimoraicity generalization, the above-mentioned 'Adjunction to Mora' (Broselow 1992, Broselow et al. 1995, Watson 1999a and 1999c) has been introduced, by which the last consonant is adjoined to the preceding syllable, without being assigned a mora of its own (see 1.3.2).

Onsets are obligatory, but complex onsets are strongly dispreferred (Angoujard 1990, Abu-Mansour 1991 and 1992, Broselow 1992, Broselow et al. 1995), at least domain-internally, even if these can actually be found at a surface level as a result of suffixation or vowel deletion (see 4.3), or in loan-words (see Holes 1995, Derwing et al. 1995:110, Watson 1999a). Of an intervocalic two-consonant sequence, therefore, the first consonant will be syllabified into the coda of one syllable, the second consonant into the onset of another syllable, even if syllabification of the two consonants into one onset would not violate Sonority Sequencing. [madrasa] 'school', for instance, is syllabified as [mad.ra.sa] rather than *[ma.dra.sa]. We can thus say that the preference for simple onsets overrides Onset Maximization in this language. In a templatic approach, the directionality of syllabification should thus be left to right. This claim, however, leads to complications, since, while all Arabic varieties display this preference for simple onsets, some dialects must have a right-to-left syllabification in order to account for other phonological processes (see 4.3, 5.3).

Complex codas are also disfavoured, except domain-finally (where, however, the last consonant is extrasyllabic) (Broselow 1992, Broselow et al. 1995), but there are no restrictions on the quality of a coda consonant.

Only in some (especially North African) dialects can consonants surface as syllable peaks, with some Moroccan varieties having syllabic obstruents at the end of the spectrum (Broselow 1992, Angoujard 1990). However, only vocoids can occupy a syllable nucleus underlyingly.

---

[39] CVCCC syllables also appear domain-finally in derived environments in dialects such as San9ani, but Watson (1999a) shows that, in this variety, the last two consonants are always [t] and [ʃ], which are interpreted as the affricate [tʃ], the whole of which is extrasyllabic.

Contrary to Italian, Arabic does not seem to require underlying or lexical syllabification, especially since it does not allow onsetless syllables. In fact, Itô's (1986, 1989) templatic procedure applies particularly well to Arabic. Other authors, however, still assume lexical syllabification, with resyllabification applying at least in some varieties (see Abu-Mansour (1991) for Makkan, Selkirk (1981b) for Cairene, Watson (1999a) for San9ani). The domain of application of resyllabification may vary from dialect to dialect; in AA, it seems to be the IPh or the PU (cf. *mbaa.riH shuf.t il.mu.dii.r u. sik.ri.tiir.tuh* 'Yesterday I saw the director and his secretary').

1.4 Stress assignment

This section will be looking at stress assignment and foot structure. Since the syllabic nucleus constitutes the stress-bearing unit, there is a strong link between accent placement and the distribution of vocoids, and processes affecting this class of sounds often interact with stress assignment in interesting ways. Moreover, foot structure, which is inextricably related to stress assignment, can also have some bearing on the application of such processes.

1.4.1 Stress

Contrary to what Chomsky and Halle (1968) seem to assume, stress is not an on-off element like voice or nasality, and therefore cannot be characterized in absolute terms by the feature [±stress]. It can only be analyzed as a relative entity, i.e. as the phonetic prominence of one syllable over the adjacent ones. Liberman and Prince (1977) first defined stress locally, making use of binary branching trees: a syllable is strong only relative to the syllable with which it is coupled, i.e. to its sister. Moreover, the notion of stress does not define a specific phonetic entity, but has several correlates, which vary cross-linguistically in both number and quality; most relevant in languages such as Italian and Arabic are loudness, pitch, vowel length, vowel quality.

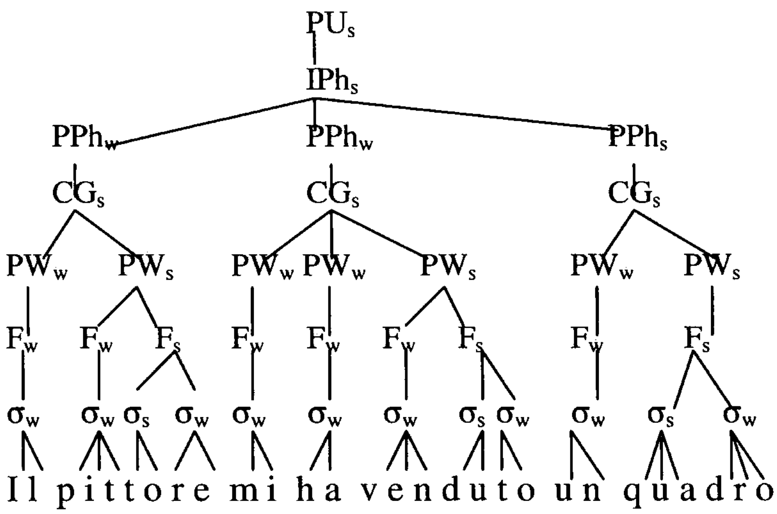
1.4.1.1 Trees and grids

1.4.1.1.1 Prosodic trees and stress assignment

In Prosodic Phonology, each phonological constituent apart from the syllable functions as a domain for stress assignment at that given level, so that inside a domain X^n there is a constituent X^{n-1} which is more prominent than all others, i.e. a

Designated Terminal Element (DTE). In other words, one node inside each domain is labelled strong (s), while all other nodes are labelled weak (w). The distribution of stressed syllables can thus be represented directly on the prosodic tree. As was said in section 1.2, contrary to Liberman and Prince's (1977), NV's trees are not binary, but n-ary branching, as in (26).

(26)



'The painter sold me a picture'

It is clear from the figure that, in this model, constituents below the CG (notably PW and F) may lack a strong node. This happens when a foot is made up of only one weak syllable, i.e. in the case of clitics (when the *degenerate foot* is the only F in the PW),[40] or in the case of a word initial weak syllable. Such an interpretation is clearly a direct consequence of the SLH. However, while unstressability is acceptable in the case of clitics because of the peculiarity of this category, the use of unstressed feet is an undesirable solution. As was pointed out in 1.2, weakening – or even complete disregard – of the SLH is to be preferred to the postulation of subminimal, degenerate units, so that many scholars now accept that material may be parsed into non-contiguous higher structure (violating Strict Succession), or even that feet may stretch

[40] In Nespor and Vogel's (1989:72) words,

> the label on non-branching nodes depends on whether the element in question is a function word or not [...] a non-branching node [is] weak iff it exhaustively dominates a clitic or one of the members of a closed set of function words.

across word-boundaries (violating Proper Bracketing). More about feet and the violation of the SLH will be said in 1.4.1.2.

The relative prominence between weak nodes in a given domain X is given by 'stratification' of strong and weak nodes in the various domains below X, so that domains from the PW up can have secondary stresses. Inside an IPh, for example, even if only one PPh is the DTE, certain PWs, feet and syllables within other PPhs may still exhibit different degrees of prominence.[41] However, according to NV (85-6) there is no prediction of relative prominence relations between the weak syllables inside one F, although in the case of long unbounded feet additional rhythmic alternations may be introduced by general principles lying at the basis of grid operations (see 1.4.1.1.2).

NV (59) maintain that a string X^n may be posited as a constituent for the only reason that it accounts for the relative prominence relations inside X^{n+1}. The lack of periodicity in the recurrence of stress peaks inside X^{n+1} and the lack of connection with given positions in syntactic constituents show that there must be a stress assigning principle 'other than an abstract alternation rule or a rule that makes direct reference to syntactic structure' (ibid.), so that another type of constituent that accounts for the stress pattern is motivated.

The principles of stress assignment vary cross-linguistically, but can generally be circumscribed by way of binary parameters (see 1.4.1.4). In particular, one is looking at whether syllable weight affects accent distribution and at whether the particular position of a syllable within a word or phrase attracts stress. Although these factors may interact in more complex ways, three fundamental typologies can be identified: 1. languages in which syllables are assigned prominence only by virtue of the position they occupy (e.g. French, Czech, Finnish, Polish, Hungarian, Turkish); 2. languages in which the position of stress is determined by syllable weight, i.e. in which heavy and superheavy syllables attract stress (see Arabic in 1.4.3); 3. languages in which stress is arbitrarily assigned, independently of syllable weight and location within the word or

[41] In Chomsky and Halle's (1968) framework, the relativity of stress is given by the cyclic assignment of numerical values to the stress-bearing units within each morpho-syntactic domain. In English, the DTE within each lexical item is first assigned primary stress (index = 1) by the *English Stress Rule* (ESR). Then a *Nuclear Stress Rule* (NSR) reassigns a 1 to the DTE within phrases or sentences, a *Compound Stress Rule* (CSR) within constructs such as *literature essay*. Each time stress is assigned, a *Stress Subordination Convention* automatically reduces by one all non-primary stresses within the domain, after which the innermost brackets are erased by a *Bracket Erasure Convention*, so that stress assignment can reapply to a larger string.

phrase, and is therefore stored in the lexicon as a property of a given lexical item (see Italian in 1.4.2). Only in the latter case does stress have a distinctive function.

1.4.1.1.2 The metrical grid

Another way of representing stress is by means of the *metrical grid*, a notational system first introduced by Liberman (1975) and developed, among others, by Halle and Vergnaud (1987). Within this approach, the metrical plane is separated from the syllable plane, and stress levels are represented by layers of asterisks located above the stress bearing units. Each syllable nucleus, i.e. each potentially stressable position, projects one asterisk on the first level (line 0). After that, each actually stressed syllable projects an asterisk on line 1, which represents the foot level. Then the DTE of each word projects a further asterisk on line 2 (the word level), and so on, with a line of asterisks for each prosodic constituent. The sentence in (26) can thus be represented as in (27).

(27)

```
                                                                          *           line 5 (PU)
                                                                          *           line 4 (IPh)
                *                                  *                      *           line 3 (PPh)
                *                  *               *                      *           line 2 (PW or CG)[42]
*               *            *     *               *          *           *           line 1 (F)
*       *       *   *        *     *       *       *   *      *           *       *   line 0 (σ)
```

Il pittore mi[43]ha venduto un[43]quadro

'The painter sold me a picture'

The debate on whether trees or metrical grids are superfluous and should therefore be eliminated from phonological representations has gone on for decades and is not yet settled. Liberman and Prince (1977) propose that grids should coexist with trees, while in Kiparsky's (1979) view only the latter are necessary; Prince (1983) and Selkirk (1984a), on the other hand, argue that the use of grids makes trees redundant. One of the disadvantages of the metrical grid with respect to prosodic trees is that, by

---

[42] The two lines are conflated because clitics do not have a stress of their own.

[43] In connected speech, *mi* and *un* do not actually form independent syllables (see Glide Formation in ch. 2), but we will assume that this is an intermediate stage in the derivation in which resyllabification (and refooting) has not yet applied.

not showing where each constituent begins and ends, it seems incapable of representing its headedness, i.e. the location of its DTE (see 1.4.1.2). However, this problem is solved by Halle and Vergnaud (1987) by building prosodic structure into the grid by means of bracketing, as in (28).

(28)

```
                            *
(*                          *)
(*)                (*       *)
(*       *)        (*   *)  (*      *)
p a i n f u l   r e c o l l e c t i o n
```

1.4.1.1.3 Clashes and lapses

The main problem with prosodic trees is instead the definition of *stress clash* – i.e. the presence of two contiguous strong elements - and *lapse* – i.e. a string of (more than two) adjacent weak elements. Although it is possible to represent these in terms of trees, the picture becomes much clearer when the metrical grid is adopted.

In fact, Nespor and Vogel (1989) analyze stress clashes – and, to a lesser extent, lapses – in terms of the metrical grid. They show that clash avoidance is a universal feature of natural languages, by which adjacency of two or more strong nodes is dispreferred,[44] but note that 'both the specification of what counts as 'adjacent' and the minimal degree of stress required for there to be sufficient pressure for change' is subject to cross-linguistic variation (ibid.:70). In Italian and English, for instance, the minimal stress clash is defined at the third – i.e. the PW – level, while in a language like Greek a clash can take place between F level stresses (ibid.:92). On the other hand, whilst in Italian and Greek, as in all syllable-timed languages[45], 'adjacency is defined starting at the lowest grid level' (ibid.:91), in English (stress-timed language)

[44] Both clashes and lapses are dispreferred on the basis of a eurhythmic principle favouring maximal alternation, which seems to underlie all Obligatory Contour Principle (OCP) effects.

[45] A language is syllable-timed when the temporal distance between two successive syllable nuclei is maintained approximately constant, while the distance between two successive stresses is not. Vice-versa, in stress-timed languages the intervals between stressed syllables are more equal than the ones between syllable peaks. To these Donegan (1985:95) adds a third class of mora-timed languages, in which there is a distinction between short (one-beat) and long (two-or-more-beat) syllables, so that each short vowel is mapped onto one beat, and each long vowel is mapped onto two beats.

two stresses at a given level *n* are adjacent, and thus clashing, when they are not separated by a stress on level *n-1*.

A number of different processes can be applied to resolve a clash, and what process will be adopted in a particular case also varies cross-linguistically. One of the most common clash-avoiding devices is so called *Stress Retraction* (SR), by which the first of the two clashing stresses shifts to the previous syllable, as in the Italian example in (29).[46]

(29)

SR:

```
           *               *
*      *   *     *    *    *
*    * *   *     *    * *  *
* *  * *   * *   * *  * *  * *
era metá stráda → era méta stráda        '(it) was half-way'
```

Another possible clash-avoiding process is the insertion of a beat between the two stresses, i.e. *Beat Insertion* (BI)[47].

[46] According to Nespor and Vogel (ibid:76), there is no actual stress shift process: the two rules of *Beat Deletion* (BD, a possible remedy to a clash) and *Beat Addition* (BA, a possible remedy to a lapse) combine together to produce such an effect.

(ii)

```
                BD:           BA:
          *             *              *
*     *   *    *        *     *   *    *
*   * *   *    *   * *  *     *   * *  *
* * * *   * *  * * * *  * *   * * * *  * *
era metà stràda → era meta stràda → era méta stràda '(it) was half-way'
```

(ibid.:76)

[47] The physical correlate of BI is either the insertion of a pause or the lengthening of the first clashing syllable. Pitch distance is another possible phonetic realisation of BI, since the basic notion is that of change (Nespor and Vogel 1989: 93; also see Lerdahl and Jackendoff 1983).

(30)

```
                              BI:
        * *                       *    *
        * *    *   *              *    *    *    *
  *  *  * *    *   *         *  *  *    *    *    *
  *  * * * * *  * *  *   *     *  * * * *  * *  * *  *    *
la veritá vínce quasi sempre → la veritá _ vínce quasi sempre
```

'truth wins almost always'

(from Nespor and Vogel 1989:78-9)

A stress lapse, on the other hand, is defined in relation to the first two levels (ibid.:83). *Beat Addition* (BA), i.e. the increase by one degree of one of the weak stresses, applies as a lapse-avoiding device, provided it does not create a clash (31).

(31)

```
    *           *          *          *
    *           *          *       *  *
 *  * * *   * *  * * *    * * * *   * *  * * *
teléfonaglielo súbito → teléfonaglièlo súbito
```

'Telephone him/ her about it immediately'

(from Nespor and Vogel 1989:85)

1.4.1.1.4 Linguistic and rhythmic stress

Nespor and Vogel (1989) thus conclude that both prosodic trees and metrical grids should be adopted, since they carry out different functions: while trees are better capable of representing prosodic structure and its relationship with stress assignment, grids can better illustrate rhythmic phenomena such as stress clashes and their resolution. In other words, prosodic trees characterize what Nespor (1993:254) terms 'linguistic stress' by showing where the strong and weak nodes occur, while the metrical grid is said to represent 'rhythmic stress'. This is constructed by mapping an abstract representation of the rhythmic structure of a string onto its prosodic structure (Nespor and Vogel 1989:70).

According to Nespor and Vogel (ibid.), linguistic and rhythmic stress belong to different cognitive domains: the former is part of the grammar, and is obligatory, while the latter is 'a linguistic manifestation of rhythm', being optional and variable, influenced by such factors as emphasis and focus relations. While the grid is built on the information provided by prosodic trees, such information can be recovered from a grid representation only so long as the rhythmic pattern of a string has not been modified. Consequently, rhythmic rules cannot affect the prosodic make-up of a string, but only its surface realization (ibid.:71).

Although both trees and grids will be assumed, for reasons of economy in this thesis accent will only be transcribed when relevant to the phenomena under analysis, and even then the tree or grid representation will be replaced by simple stress marks: ⟨ˊ⟩ for primary (word) stress, ⟨ˋ⟩ for secondary stress.

1.4.1.2 The foot

The minimal domain for stress assignment is the foot, which can actually be defined as a constituent formed by a stressed syllable plus all the unstressed syllables to its right or left until the next stressed syllable. Feet can in fact be *right-headed* (or *iambic*) or *left-headed* (or *trochaic*), according to whether the strong node is rightmost or leftmost. Moreover, feet can be either *bounded* or *unbounded*. Bounded feet are generally disyllabic, although the existence of trisyllabic bounded feet has often been acknowledged (see NV, Halle and Vergnaud 1987, Idsardi 1992, Halle and Idsardi 1995, Roca and Al-Ageli 1999). Other scholars (e.g. McCarthy and Prince 1986, 1990, Prince and Smolensky 1993, Burzio 1994, Hayes 1995), only accept binary feet. Such binarity, however, can be defined either at the level of the syllable or at that of the mora, so that monosyllabic bimoraic feet are generally allowed. Monomoraic (also called 'degenerate') feet, although permitted by NV in order to satisfy the SLH, are instead disallowed on the basis of minimality requirements. Moreover, such feet would be stressless, thus contradicting the selfsame definition of foot.[48] Unbounded feet, on the other hand, may in principle contain any number of syllables.

A third parameter along which languages vary is whether or not syllable structure can determine foot structure, i.e. whether syllable weight can determine stress placement (see 1.4.1.1 above). According to this parameter, feet are classified as *quantity-sensitive* or *quantity-insensitive*.

---

[48] Some languages, however, exhibit stressed degenerate feet. As will be shown in 1.4.2, such constituents can actually be seen as bimoraic feet with an empty position.

Just as segments must be licensed by syllables, syllables must be licensed by feet. Nonetheless, for scholars who do not accept ternary feet - and also for those, i.e. the majority, who do not accept monosyllabic unstressed feet - some syllables must necessarily remain unparsed (*partial foot licensing*). There is, however, a strict limit of at most one unparsed syllable in a row (Hammond 1997:43). Moreover, as was mentioned above, unparsed syllables are always unstressed: a single stressed syllable does form a foot (ibid.). Since such unfooted syllables are generally phonetically realized, they must be incorporated into prosodic structure at a later stage, directly adjoining at the level of some higher constituent such as the PW. They may thus be seen as extrametrical (cf. Hayes 1980).[49] Extrametricality, however, does not apply to monosyllables, according to a *Non-Exhaustivity* condition: 'an extrametricality rule is blocked if it would render the entire domain of the stress rules extrametrical.' (Hayes 1995:58). Moreover, Hayes (1995:107) formulates a constraint against *chained extrametricality*: 'Extrametricality does not chain; i.e. a constituent followed by an extrametrical constituent is not counted as peripheral'. This is why no more than one syllable in a row can go unparsed.

According to Brame (1970, 1972, 1974), a phonological string is reparsed into feet at the end of every lexical cycle. Postlexical processes like syncope and epenthesis, on the other hand, apply post-stress assignment (see ch. 4 and 5). However, at a surface level feet often appear to extend across PWs, thus violating the Proper Bracketing condition of the SLH. This brings up the question of whether these constituents may belong to a different hierarchy or even be analyzed as mere rhythmic units with no phonological counterpart.

The problem of the particular status of feet in the prosodic hierarchy has been tackled in different works, among which a paper by Eric Fudge. In his article, Fudge (1999) challenges the assumption that PW, F and σ fall into a single phonological hierarchy, and in particular that feet are always sub-units of PWs, based on the fact that word and stress group boundaries may not coincide with foot boundaries.[50] His model has two intersecting hierarchies, with PU, IPh and σ belonging to both, and PW, CG and

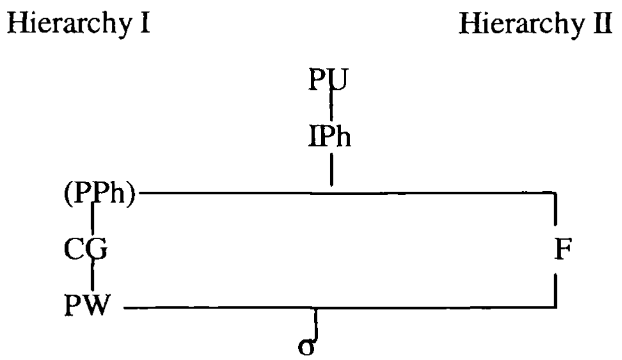
---

[49] Scholars who do not accept ternary feet, and thus allow for word-internal unfooted syllables (cf. Hammond's (1997) *[Winni]pe[sáukee]*), cannot invoke extrametricality, and thus encounter difficulties justifying the phonetic realization of such syllables.

[50] Fudge (1999) also finds a redundancy in the number of prosodic constituents NV postulate: some of them, like IPh and PU, or PW and CG, form a closely related pair.

perhaps PPh - whose existence as a prosodic constituent Fudge (ibid.:279) actually questions - belonging to a different hierarchy than F (cf. fig. (32) below).

(32)



Other proposals exclude the F from the prosodic hierarchy altogether. Since the notion of F is strictly related to that of accent placement, it is not surprising that most of the arguments against - as well as in favour of - the inclusion of this constituent in the prosodic hierarchy are based on stress assignment. Scholars like Prince (1983) and Selkirk (1984a) point out that, if the F is defined on the basis of stress assignment, and if stress patterns can be characterized in terms of the metrical grid, then it is not economical to keep both representations in the same model. The above-mentioned proposal of Nespor and Vogel's (1989) to keep both trees and grids in order to characterize two different kinds of stress does not necessarily clear the doubts about the existence of feet. Nespor (1993:170), on the other hand, argues that the fact that stress distribution conditions the application of a rule can be expressed in terms of F constituents, without resorting to the grid at a linguistic level.

Although the debate is still open, for the purposes of this thesis I will assume that feet exist as phonological units and that they are part of the prosodic hierarchy. As will become clearer in the following chapters, it is often useful to refer to F structure when characterizing the distribution of vocoids.

1.4.1.3 Word stress

In spite of what was said above, the stress that most affects the application of phonological rules, even if these apply in larger domains, is PW-stress. This appears to be more prominent and more readily perceivable than stress in other domains, and has thus been denominated *main stress*.

The characteristics of PW-stress vary cross-linguistically. In some languages, it has a distinctive value – i.e. two separate lexical items can be distinguished merely on the basis of stress placement – while in others it does not. Besides, as was mentioned in 1.4.1.1.1, in some languages stress is fixed – i.e. it always falls on the same syllable within each word – while in other languages it is mobile. This variation can be accounted for on the basis of the fact that relative prominence inside the PW depends on two parameters: (1) Whether the first or the last F in the PW is stressed, and (2) whether feet are left-headed (i.e. trochaic) or right-headed (i.e. iambic). The combination of these two parameters yields four possibilities:

a. Right-headed PW, right-headed F = stress always on the last syllable.
b. Left-headed PW, left-headed F = stress always on the first syllable.
c. Right-headed PW, left-headed F = stress on one of the last three syllables, according to whether the F is trisyllabic, disyllabic or monosyllabic.
d. Left-headed PW, right-headed F = stress on one of the first three syllables, according to the above factors.

1.4.1.4 Stress and syllable structure

As was mentioned above and in 1.3, the distribution of stress, and especially PW-stress, is often connected to syllable structure. The correlation is bidirectional, in that particular types of syllable do or do not attract stress, but at the same time the presence or absence of stress on - or even its position outside of - a given syllable can determine changes in its structure. This is not surprising, given that the syllable nucleus is the stress bearing unit: the minimal domain in which stress is assigned is the foot, of which syllables are the immediate constituents. Moreover, accent is related to the sonority and length of segments, which are also determining factors in syllabification. For example, vowel length may[51] determine syllable weight, and in many languages heavy syllables attract stress. On the other hand, one of the physical correlates of stress is vowel length, so that a stressed vowel is – phonetically, and sometimes phonologically – lengthened. Prosodic phenomena which are connected with syllabification and resyllabification are easily affected by stress placement, and

[51] If a syllable has two – or, in many languages, even one - coda consonants, the length of its nuclear vowel does not make a difference from the point of view of stress assignment.

vice-versa. Vowel reduction processes such as Syncope, Shortening or Glide Formation often appear to be blocked by the presence of stress on the target segment. Not only vowel quantity, but also vowel quality can be determined by accent placement. Segmental quality is actually another physical correlate of stress, so that, according to whether it bears prominence or not, a segment can be strengthened or weakened respectively, thus changing its position on the sonority scale. In English, for instance, unstressed vowels are generally reduced to [ə] (e.g. *narration* [nə'ɹɛɪ̯ʃən]). By the same token, the Italian vowels /ɛ/ and /ɔ/ are raised to [e] and [o] respectively when they do not bear primary word prominence (cf. NV:130-1).[52]

When dealing with features such as vowel length, which do not involve resyllabification but only syllable internal alternations, the question may arise of whether it is syllable structure that affects stress assignment or stress assignment that affects syllable structure. As stated above, for many languages it has been assumed that syllable weight, and thus vowel length, determines the position of stress. In other languages, however, it is clearly stress placement that determines vowel length (see Italian in 1.4.2 and 3.2).[53]

1.4.2 Stress assignment in Italian

Italian exhibits a mobile stress accent that has a distinctive function, with a wealth of minimal pairs such as *príncipi* 'princes'/ *princípi* 'principles' or *fáro* 'lighthouse'/ *faró* '(I) will do'. According to Nespor (1993), feet in Italian are bounded and trochaic – i.e. the first syllable in a foot is its DTE - and can be either binary, ternary or unary (degenerate). Since PW stress falls on the last foot in a word, a combination of these two parameter settings will give rise to the following possibilities:

a. Stress on the penultimate syllable: *re[císo]$_F$* 'cut off (past part.)'
b. Stress on the ante-penult syllable: *[récito]$_F$* '(I) act'
c. Stress on the last syllable: *reci[tó]$_F$* '(s/he) acted'

---

[52] The quality of non-nuclear segments can also be affected by the presence or absence of an adjacent stress. This is the case of Voiceless Stop Aspiration in English, which tends to be blocked in unstressed syllables (e.g. *atom* ['ætəm] vs. *atomic* [ə'tʰɒmɪk]).

[53] The same question may be asked as regards the segmental level: does the distribution of stress determine segmental quality and quantity, or is it the latter which determines stress patterns? Based on the idea that it is actually segmental quality and quantity that determine stress patterns, Nespor (1993:258-9) suggests that the distinction between syllable-timed and stress-timed languages (see fn. 45) should be eliminated.

It is impossible to give a rule for PW stress assignment in Italian without referring to the notion of foot. Even if we allow for long vowels to be present underlyingly (but see 3.2), representing a word like *recíso* as /retʃi:zo/, there is no way to distinguish between the two forms *récito* and *recitó* based on the idea that heavy syllables attract stress, since the root *recit-* is the same and should always have the same UR.

There is clearly a strict correlation between main stress and bimoraicity in Italian, where stressed syllables are either of the shape CVC or CVV at a surface level. As will be shown in section 3.2, accented vowels are always lengthened in open syllables.[54] A striking exception to such a pattern are PW-final stressed open syllables, which in this language are monomoraic. Based on phenomena such a Syntactic Doubling, which copies a PW-initial consonant into the coda of a preceding stressed open syllable (e.g. *caffè nero* 'black coffee' > [kaf.fɛ**n.ne**:ro]), it has been claimed (Selkirk 1984a, NV) that such syllables are in fact bimoraic underlyingly, with an empty coda position which is then filled by the reduplicated consonant.

1.4.3 Stress assignment in Arabic

Arabic is a syllable-timed – or better, according to Donegan's (1985) definition (see fn. 45), mora-timed - stress accent language. Stress is mobile but it is not distinctive. Extremely rare minimal pairs have been claimed for some dialects, but according to Holes (1995) their existence is debatable. This is related to the fact that there is an underlying distinction between long and short vowels even in unstressed syllables, and that relative prominence is determined by syllable structure and by the overall shape of the PW.

As in most, perhaps all, Arabic varieties, in AA feet are trochaic, and stress falls on the last F. However, weight considerations are foremost. Although there is a certain

[54] Phonetic length, however, is conditioned by foot structure, too: the stressed vowel in a three-syllable word (and foot) like *cómico* 'comic, comedian' is shorter than the one in a disyllabic word (and foot) like *cóme* 'how; like, as'. Again, these considerations raise the question of whether there is a clear-cut distinction between syllable-timed and stress-timed languages.

cross-dialectal variation,[55] general rules for main stress assignment can be formulated, and are followed rather strictly by Levantine varieties like AA. These are as follows:[56]

1. Stress a superheavy final syllable (e.g. *tadxuláan* '(you two) enter')
2. Otherwise, stress the rightmost heavy syllable up to the penult (e.g. *daxálna* '(we) entered')
3. Otherwise, stress the antepenult (e.g. *dáxala* '(he) entered')

(Mc Carthy 1979a, Abu-Mansour 1991 and 1992, Broselow 1992, Majdi and Winston 1994, Holes 1995, Younes 1995, Watson 1999a)

We can thus say that, as in Italian, PW stress falls on the rightmost foot, that feet are trochaic and minimally bimoraic, and that their structure is affected by syllable weight and by the structure of the word.

In Classical Arabic, a word-final open syllable containing a long vowel is also stressed. In other varieties, however, both CVV and CVC syllables in final position fail to be accented (see for instance Makkan Arabic in Abu-Mansour 1992:52). While CVC can be considered light on the basis of the extrasyllabicity of the final consonant, CVV would appear to be heavy. Lack of prominence in both cases can be accounted for by postulating that the whole final syllable is extrametrical, unless it is followed by an extrasyllabic consonant, as in CVVC or CVCC. In AA, although all word-final vowels surface as short (cf. 3.3), this generalization can be maintained as regards vowels that were underlyingly long.

In some Arabic dialects, vowel sonority sometimes appears to play a role in stress assignment, too. For instance, according to Majdi and Winston (1994:200), in Cairene Arabic stress cannot be assigned to an open syllable with a [+high] vowel.

[55] In San9ani Arabic, for example, CVC syllables are only stressed when in the environment of other CVC or CV syllables, but not when in the environment of CVVC, CVV or CVG (G = geminate) syllables. In Watson's (1999b) two-layered model (cf. 1.3), foot construction first looks at the upper level, considering the lower layer - where CVC syllables are bimoraic - only if bimoraic syllables are not available at the upper one.

[56] The examples are from Standard Arabic.

1.5 Optimality Theory

OT is a non-derivational approach which was first proposed by Prince and Smolensky (1993, henceforth PS). Since it has been predominantly applied to phonology, which is also the main interest of this thesis, the present treatment will be limited to such a field. In section 1.5.1 the tools of OT will be introduced, looking first at the notions of universality, violability and constraint ranking, then at markedness and faithfulness constraints, at constraint conflicts and candidate evaluation, and finally at a comparison between OT constraints and DT rules. Section 1.5.2 will deal with prosodic structure constraints and their interaction, focusing in particular on the syllable, the foot and stress assignment. 1.5.3 will center on morpho-phonological constraints, while 1.5.4 will briefly touch upon output-to-output correspondence.

1.5.1 A theory of constraints

1.5.1.1 Universality and violability

In PS's (2) words, OT is 'a theory of the way that representational well-formedness determines the assignment of grammatical structure'. It is based on the assumption that linguistic patterns can be accounted for on the basis of a set of universal constraints, doing away with derivational rules.

Clearly, the concept of a universal foundation to human language is already present in DT. Ever since early Generative Theory (Chomsky 1957, 1965, Chomsky and Halle 1968), basic traits common to all languages have been identified, and described as Universal Grammar (UG). However, this is restricted to a limited set of properties, especially in early Generative works (Chomsky 1965, Chomsky and Halle 1968), where the overwhelming majority of alternations are accounted for by means of language-specific *rewrite rules*. In the later 1970's-early 1980's, more and more universal conditions were introduced (e.g. the OCP in phonology), in an attempt to constrain the form and typology of rules. This led to the formulation of a Principles-and-Parameters Theory (Chomsky 1981, Hayes 1980), which purports that the grammar of individual languages is based on the combination of core universal properties (*principles*) and binary choices (*parameters*). Besides, models such as Underspecification Theory and Lexical Phonology were designed to constrain rule application and rule interaction (see Kager 1999:2 for further discussion).

The common assumption of all the above-mentioned theories is that universal principles are absolute, and cannot be violated. Ever since Structuralism, and all through Generative Phonology, however, a less strict notion of universality has been adopted (especially in linguistic typology): *markedness*. This is the concept that all forms of language structure have two values, one of which is cross-linguistically preferred over the other. The preferred value is said to be *unmarked*, the dispreferred one *marked*. As Kager (ibid.) points out, markedness is a relative notion: an element is not ill-formed per se, but only in comparison to other elements.

Markedness is a key concept in OT. This theory still accepts the notion of UG and, at least in its standard form, the idea of an underlying representation (the input form); moreover, it makes use of distinctive features and sometimes even of prosodic and morphological tree diagrams.[57] However, it is based on a radically different conception of grammatical alternation with respect to standard DT.

Instead of using rules that modify the input of phonological processes, it focuses on the output forms, on which it places constraints. Such constraints, contrary to derivational rules, are *universal*, i.e. they apply in all languages (they are all part of UG).[58] However, they are *violable*. In fact, constraints are naturally in continuous conflict with one another, i.e., in order to satisfy one constraint, another constraint must be violated. From this follows that no output candidate, not even the optimal one, can satisfy all constraints (*Fallacy of Perfection*): 'what determines the best output of a grammar is the least costly violation' (Kager 1999:3).

In fact, constraints are ranked so that violation of a lower-ranked constraint *B* is less serious than, and thus preferred over, violation of a higher-ranked constraint *A*. In this case, one can say that *A* dominates *B*, i.e. *A* » *B*. One property of constraint ranking is transitivity, so that, if *A* » *B* and *B* » *C*, then *A* » *C*. Languages differ precisely in the way they rank the universal output constraints, and thus in the way they resolve

[57] However, OT has always exhibited a tendency towards the reduction of representations in favour of constraint interaction. Some authors have actually aimed at eliminating URs by introducing allomorphic models (cf. Kager 1999:413-20), and arguments have been put forward against the use of underspecification (Inkelas 1995, Itô, Mester and Padgett 1995, Steriade 1995), feature geometry (Padgett 1995) and even syllable structure (Steriade 1995).

[58] In its strongest formulation, OT requires all constraints to be universal, i.e. to be present in the grammar of all languages. It appears, however, that the universality of some constraints should be partly relativized, in the sense that language-specific elements should be introduced into constraints of otherwise universal format. This is especially the case with interface constraints, such as those of the ALIGNMENT family (see 1.5.2.2.1), which supply universal schemata in which individual languages may insert specific morphemes.

conflicts between them. Consequently, a constraint may enforce an absolute ban on some phonological feature in one language (where it is undominated), a preference in another language (where it is dominated by some constraints) and have no effect at all in a third one (where it is ranked so low that it is virtually rendered inactive).

A typical example of constraint interaction is that between the two potentially contrastive constraints *CODA – by which empty syllable codas are preferred – and *COMPLEX – by which complex onsets are disallowed. In a situation in which two consecutive consonants have to be syllabified, a clash takes place, and one of the two constraints must be violated.[59] If, in a given language, *CODA » *COMPLEX, the latter will be violated, yielding a complex onset (V.CCV). In a language in which *COMPLEX outranks *CODA, on the other hand, the first consonant will be syllabified into the coda of the first syllable (VC.CV).[60]

The degree of satisfaction of the constraint hierarchy is called *harmony*. One says that a candidate *a* is more harmonic than a candidate *b* (*a* ≻ *b*) when *a* incurs less serious violations than *b*, and is therefore preferred.

1.5.1.2 Constraint types

The two main types of constraints are *markedness constraints* and *faithfulness constraints*. The former aim at reaching structural well-formedness by eliminating

---

[59] Of course, other possibilities arise. A possible scenario is that in which both the above mentioned constraints outrank a third constraint called FILL, by which an empty position must be filled. In this case, only FILL will be violated, yielding the insertion of an epenthetic vowel between the two consonants (see ch. 5; also see PS:25-6).

[60] These considerations clearly have fundamental implications for language acquisition: acquiring a language, in an OT perspective, means acquiring a lexicon plus a language-specific hierarchy of universal constraints. This must be reconstructed starting from the output forms, which are the only ones the learner has access to, i.e. the constraints must be ranked in a way that predicts all and only the output forms of that specific language. In order to account for this, Tesar and Smolensky (1993) devise a *constraint demotion algorithm*, by which information on constraint ranking is extracted from constraint violations, based on the notion that constraints which are violated in the optimal candidate must be dominated by constraints which are unviolated in the optimal candidate but violated in some suboptimal candidate. The algorithm consists in a step-by-step reranking of the universal constraints based on the 'demotion' of a constraint to a lower position in the hierarchy every time this is shown to be dominated by another constraint (*Recursive Ranking*). Demotion must be minimal, i.e. a constraint is demoted to a position immediately below the highest-ranking constraint that induces its violation in the optimal candidate. At every step, sets of constraints will be formed which are still unranked with respect to one another: these are called *strata*; a hierarchy that contains one or more strata is called a *stratified* hierarchy, and is a temporary hypothetical hierarchy that arises in the course of the algorithm, representing the current knowledge of the learner. As the learner is presented with new output forms, the hierarchy is refined, though never completely subverted, until no additional rerankings are motivated on the basis of the positive evidence available to the learner (ideally, until the hierarchy is no longer stratified).

marked forms. Examples of markedness constraints are *[+round] (no vowels are [+round]), *VOICED-CODA (coda consonants are voiceless), PARSE-SYL (all syllables must be parsed). As is evident from the examples, they may either take the form of prohibitions of marked phonological structures or of positive requirements of unmarked structures. Markedness constraints are blind to the lexical input (they are purely output constraints), and their universality is based on two criteria: typological distribution and phonetic grounding.[61] There are two types of markedness constraints: a. *context-free* markedness constraints, which express a prohibition or a requirement as an absolute, in any context (e.g. $*V_{NASAL}$ = 'vowels must not be nasal'); b. *contextual* markedness constraints, which express a prohibition or requirement in relation to a specific environment (e.g. $*V_{ORAL}N]_{\sigma}$ = 'before a tautosyllabic nasal, vowels must not be oral').

Faithfulness constraints preserve lexical contrasts by imposing analogies between output forms and the correspondent input forms. Examples of faithfulness constraints are IDENT-IO(voice) ('correspondent consonants in the input and the output are identical in their specification for voice'), LINEARITY-IO ('the output reflects the precedence structure of the input, and vice-versa'), MAX-C-IO ('input consonants must have output correspondents'). Faithfulness constraints are not pure output constraints, since they also refer to the input. However, contrary to DT, OT postulates no constraints that concern exclusively the input.

Faithfulness constraints, apart from preserving lexical contrasts, also put 'pressure towards the *shape invariability* of lexically related items in various grammatical contexts', thus preserving similarities between different alternants of a single morpheme (Kager 1999:6). Strictly related to the concept of faithfulness is that of *correspondence*, i.e. the notion that each input element has a particular 'realization' (i.e. a *correspondent*) in the output. *Correspondence Theory* is actually a subtheory of faithfulness constraints, first proposed by McCarthy and Prince (1995).[62] It makes use of correspondence constraints, which ban divergences between the input and the

[61] For a constraint to be absolutely valid, both criteria should be satisfied. This, however, is not the case for many of the constraints that have been proposed within the OT framework. In particular, the typological criterion, in itself circular, should always be validated by phonetic grounding, but for many OT constraints this latter has not been proven as yet.

[62] *Correspondence Theory* replaces PS's assumption that no element may be removed from the input form, so that the whole input is contained in each candidate output, even when it is left unrealized (*Containment Theory*).

output with respect to some specific property. Typical correspondence constraints are for example IDENT-IO(X), requiring that an output element be identical to the corresponding input element for some feature X, UNIFORMITY-IO, requiring that no element of the output have multiple correspondents in the input ('No coalescence'), or MAX-IO, requiring that input segments have output correspondents ('No deletion', see ch. 4). In fact, the family of correspondence constraints can be divided into sub-families such as IDENTITY, LINEARITY, MAXIMALITY, DEPENDENCE, UNIFORMITY, CONTIGUITY or ANCHORING.

1.5.1.3 Conflicts

Markedness constraints and faithfulness constraints are inherently in conflict, since the latter preserve oppositions, and in every opposition one member must be marked. Since languages which accord absolute priority to faithfulness or to markedness do not exist, within one single language faithfulness prevails over markedness in some (context- and/or feature-specific) conflicts, while markedness prevails on faithfulness in other conflicts.

Domination of a markedness constraint over a faithfulness constraint results in neutralization of contrast, while domination of a faithfulness constraint over a markedness constraint results in preservation of contrastivity. Taking into consideration the distinction between contextual and context-free markedness, Kager (1999:34-43) outlines a factorial typology of the interaction of markedness and faithfulness, which can be summarized as in (33).

(33) a. *context-free markedness* » *contextual markedness, faithfulness* = lack of variation (the unmarked value always surfaces)

b. *contextual markedness* » *context-free markedness* » *faithfulness* = allophonic variation

c. *contextual markedness* » *faithfulness* » *context-free markedness* = positional neutralization

d. *faithfulness* » *contextual markedness, context-free markedness* = full contrast

As was shown in 1.5.1.1 with the *CODA - *COMPLEX example, conflicts between two constraints of the same type (in this case context-free markedness) may also take place.

1.5.1.4 Evaluation of constraints

As was mentioned in 1.5.1.1, although the grammar generates an infinite set of output candidates,[63] only one is selected as optimal (i.e. most harmonic), on the basis of its incurring the least costly violations with respect to the language-specific constraint hierarchy. Parallel evaluation of all candidates is performed by sifting them through the hierarchically ranked constraints, from the highest to the lowest ranked: each constraint may eliminate one or more of the candidates, until only one is left, i.e. the optimal output. In fact, as was mentioned above, a lower ranked constraint may be violated to avoid the violation of a higher-ranked constraint, though violation must always be kept to a minimum, as the economy property in (34) predicts.

(34) **Economy**

> Banned options are available only to avoid violations of higher-ranked constraints and can only be used *minimally*.
>
> (PS:27)

Evaluation also follows the principle of *strict domination*,[64] by which 'violation of higher-ranked constraints cannot be compensated by satisfaction of lower-ranked constraints' (Kager 1999:22).

An OT grammar thus needs essentially three components: (1) a *Lexicon*, which contains the input forms, i.e. the underlying forms of morphemes, (2) a component that maps an input form onto an infinite set of output candidates, called the *Generator*, or *Gen*, and (3) a component that evaluates the candidate output forms by

---

[63] Kager (1999:21), however, points out that excessively unfaithful candidates must be automatically ruled out, since they violate faithfulness without compensation from reduction in markedness.

[64] Strict domination also entails the assumption that, in each language, all constraints are ranked with respect to one another, even when lack of interaction prevents one from determining such a ranking. However, in cases of free variation – i.e. when, for a single input, there are two optimal output forms – two (or more) constraints must be allowed to stand in a relation of *free ranking*. According to Kager (1999:406), this implies that

> Evaluation of the candidate set is split into two subhierarchies, each of which selects an optimal output. One subhierarchy has $C_1 \gg C_2$, and the other $C_2 \gg C_1$.
>
> [where C_1 and C_2 are the two free-ranked constraints]

means of a set of ranked constraints, selecting the optimal output, called *(Harmonic) Evaluator*, or *(H-) Eval.*

As was stated above, markedness constraints never refer to the lexical input; this is because no restrictions hold at the level of the underlying forms (*Richness of the Base*). OT thus eliminates the use of Morpheme Structure Constraints, which have the disadvantage of duplicating information which is independently given by rewrite rules (Kager 1999:20).[65] Although underspecification is not used in OT,[66] the constraint ranking must be such that assuming either value for an input feature still yields the correct output: since markedness constraints act exclusively on the output, whether the input assumes one or the other form should be of little interest. However, in order to minimize the violations of faithfulness, in the absence of empirical evidence for one or the other input, the input should be selected that requires the minimum number of violations to yield the given output form (*Lexicon Optimization*, PS:192).[67]

An important principle concerning *Gen* is so-called *Freedom of Analysis*. This asserts that, for a given input, *Gen* is free to generate any conceivable output candidate that is in accordance with UG (i.e. that is made of possible linguistic structures).[68]

As for *Eval*, it is the most important component, containing the set of all universal constraints *Con*, the means to assess violation marks, to rank an infinite set of candidate output forms according to their degree of harmony and to select the optimal candidate.

---

[65] This is the so-called *Duplication Problem* (Kenstowicz and Kisseberth 1977), by which DT rules, contrary to OT constraints, fail to account for the fact that the 'dynamic' phonology of a language (i.e. the structural changes that affect lexical items) is closely related to its 'static' phonology (i.e. the structural conditions holding for all lexical items).

[66] In specific circumstances, however, a certain degree of underspecification is often assumed even in OT-based works, at least at the representational level. This is the case, for instance, of the archiphoneme N used in the treatment of Nasal Assimilation.

[67] The issue of the input forms is one of the most problematic and obscure areas within OT, precisely because of the potential contrast between the Richness of the Base, by which the form of the input appears to be irrelevant, and the need, typically associated with language acquisition, to determine the shape of the input forms in order to build a lexicon, leading to Lexicon Optimization.

[68] The fact that *Gen* generates an infinite number of candidates for each input form does not seem to constitute an unsolvable problem from a computational point of view: the evaluating algorithm always allows us to produce the correct output, whatever the number of candidates (see PS, ch.10). Also, the maximal number of violations that need to be taken into consideration is small, and computational strategies may even eliminate suboptimal candidates by classes, in the sense that candidates which share with another candidate a set of violation marks but have at least one additional violation may be eliminated automatically (Kager 1999:26).

The selection of the optimal candidate can be characterized by means of a *tableau*, i.e. a table that lists the output candidates vertically in random order, and the constraints horizontally, from the highest to the lowest ranked. A maximally simplified schema for simple domination, containing only two candidates and two constraints (with *Constr1* » *Constr2*), is given in (35).

(35)

	Constr1	Constr2
a. ☞ Cand A		*
b. Cand B	*!	

In this case, *Cand A* is the optimal candidate (marked by the index '☞'), and therefore the selected output, because it incurs the less serious violation, i.e. it satisfies the dominant constraint, which is instead violated by *Cand B* (i.e. *Cand A* ≻ *Cand B*). In the tabloid, each asterisk represents a constraint violation, while the exclamation mark indicates a fatal violation, i.e. one that eliminates a candidate. The cells following the elimination of the last candidate (in this case, the only other candidate, *Cand B*) are shaded because they are no longer relevant to the selection of the output.

A form may satisfy a constraint vacuously, when the constraint makes a requirement about some element that is not present in the candidate form.

Returning to our example of the *CODA/ *COMPLEX interaction, in a language like English *CODA » *COMPLEX, so that, when dealing with the syllabification of the input /eɪprən/, the optimal candidate will be [eɪ.prən], and not *[eɪp.rən] (see tableau in (36)).

(36)

Input: /eɪprən/	**CODA*	**COMPLEX*
a. ☞ *eɪ.prən*		*
b. *eɪp.rən*	*!	

A candidate violating a constraint more than once is provided with as many violation marks as it has violations for that constraint. Following the principle of *strict domination*, however, it is not so much the number of violations that counts, but the ranking of the violated constraints. This is shown in the tableau below: although *Cand A* exhibits multiple violations for *Constr2*, and *Cand B* exhibits only one violation for *Constr1*, the former is still selected as the optimal candidate because *Constr1* » *Constr2*.

(37)

	Constr1	Constr2
a. ☞ Cand A		**
b. Cand B	*!	

The same applies in the case of violations of multiple lower-ranked constraints, as in (38).

(38)

	Constr1	Constr2	Constr3
a. ☞ Cand A		*	*
b. Cand B	*!		

Since at least one candidate must be selected at each step of the evaluation, elimination of all candidates in the set is not allowed. Therefore, when, at some point during the selection process, all remaining candidates violate the highest-ranked remaining constraint, the seriousness of the violation becomes crucial: the forms with the fewest violation marks are selected.

(39)

	Constr1	Constr2
a. ☞ Cand A	*	*
b. Cand B	**!	

If there is still a tie between two or more candidates all having the same number of violations for the relevant constraint, then they are all passed on to the next constraint for further evaluation.

(40)

	Constr1	Constr2
a. ☞ Cand A	*	
b. Cand B	*	*!

Evaluation of the set of candidates with respect to a given constraint may be absolute or gradient. Absolute constraints are either satisfied or violated, and do not take into account the degree of violation (although a single candidate may display multiple violations of such a constraint). Gradient constraints, on the contrary, take into account the degree of violation (see ALIGNMENT constraints in 1.5.2.2.1).

1.5.1.5 Phonological processes, rules and constraints

Like DT, standard OT construes phonological processes as the mapping of an underlying representation (the *input*) to a surface form (the *output*). However, in doing so OT grammar does not make use of rewrite rules: all changes apply in one step (Kager 1999:20). In DT, application of the rules follows a *serial* order,[69] with intermediate stages in the derivation between the UR (the input) and the surface form

[69] Kager (1999:74) points out that, since the notion of faithfulness is absent in DT, prediction of economy is achieved through rule ordering, rather than being entrenched in the theory.

(the output); in OT evaluation of all the output candidates applies *in parallel*, and the mapping of the input onto the output is direct, without any intermediate steps. Although this may seem desirable in order to reduce the power of the theory, there are cases in which the outcome of a phonological change cannot be explained without referring to an intermediate level of representation, and these constitute perhaps the most challenging problem for OT.[70]

Since there are various ways to obtain a certain output, a single markedness constraint may trigger a variety of structural changes, depending on its interaction with other constraints. In particular, the repair strategy by which a high-ranking markedness constraint is satisfied depends on which of various faithfulness constraints is ranked lowest. For instance, Kager (1999:59-88) shows how, according to the language-specific constraint hierarchy, a highly-ranked *NC̥ ('no nasal plus voiceless obstruent sequences') constraint can be satisfied by means of either nasal substitution (when *NC̥ » LINEARITY-IO), or nasal deletion (when *NC̥ » MAX-IO), or vowel epenthesis (when *NC̥ » DEP-IO), or post-nasal voicing (when *NC̥ » IDENT-IO(ObsVce)), or denasalisation (when *NC̥ » IDENT-I→O(nasal)), with only vowel epenthesis being unattested. According to Kager (1999:56), 'rule-based theory fails to make [the] prediction of the functional unity of processes because it has no formal means of expressing the notion of 'output goal' of a phonological rule', or to circumscribe the specific class of repair strategies which apply in order to avoid a certain output feature.[71]

Kager (1999:53) points out that every part of a traditional rewrite rule has a correspondent in OT. Apart from the input (corresponding to the UR) and the output (corresponding to the surface form), the trigger, which in DT is the structural description, is a negative constraint, while the structural change is one out of an infinite set of changes proposed by *Gen*. In fact, under Freedom of Analysis, *Gen* may

---

[70] The phenomenon by which output alternations are based on generalizations that are not surface-true, so that reference to some intermediate level of representation seems to be required, is called *opacity*. Although various proposals – among which the introduction of intermediate levels (see Kiparsky 2000), of input-output markedness constraints, of faithfulness to other candidates (i.e. *Sympathy*, see McCarthy 1998) - have been put forward in order to solve this problem, no satisfactory solution has as yet been found within the boundaries of OT.

[71] In DT, the implicit tendency of separate rules to contribute towards a single output is called a *conspiracy* (Kisseberth 1970).

propose any kind of change, including no change at all: the outcome of the 'process' depends entirely on the selection that *Eval* makes by means of the constraint hierarchy. Obviously, if some change does take place, at least one faithfulness constraint (the lowest-ranked) has been violated.

1.5.2 Prosodic structure in OT

1.5.2.1 Syllable structure

Many OT markedness constraints on syllable structure reflect those that were introduced in DT. The Onset Principle (see 1.3), for instance, is translated into OT terms as ONSET (*$[_{\sigma}V$ = 'Syllables must have onsets'). Languages in which ONSET is undominated have obligatory onsets. That this constraint is active (although dominated by some faithfulness constraint) even in those languages in which onsetless syllables are allowed is shown by the universal syllabification of a single intervocalic consonant as an onset, rather than a coda. A clearly needed constraint is also NUC ('Syllables must have nuclei'), while *M/V ('V may not associate to Margin nodes (i.e. onset/ coda)') and *P/C ('C may not associate to Peak nodes (nucleus)') are associational constraints formulated by PS (85-7).

Another cross-linguistic generalization is that open syllables appear to be preferred over closed syllables, since no language that has closed syllables lacks open syllables. Hence, the syllable well-formedness constraint *CODA (*$C]_{\sigma}$ = 'Syllables must not have codas'), also called NOCODA or -COD. Languages in which *CODA is undominated have only open syllables.[72] Kager (1999:94) points out that this constraint is grounded in phonetic perception: coda consonants tend to be unreleased, and thus lack perceptual cues. The combined action of ONSET and *CODA makes the correct prediction that the unmarked syllable type is CV.

Some languages, as was said in 1.3, do not ban syllable codas completely, but restrict the type of segments that may appear in such a position. In particular, many languages

---

[72] PS (93) formulate a Coda Theorem:

> Codas are optional in a language if –COD [i.e. *CODA] is dominated by both PARSE and $FILL^{NUC}$.
> Otherwise, codas are forbidden.
> In the latter case, -COD is enforced by underparsing (phonetic deletion) if PARSE is the lowest ranking of the three constraints; and by overparsing (epenthesis) if $FILL^{NUC}$ is the lowest (cf. ch. 4 and 5).

exhibit a Coda Filter (Itô 1989) disallowing any coda consonant that does not share its place of articulation with the following onset consonant. In OT terms, such a filter is called CODA-CONDITION, and is formulated as in (41).

(41) **CODA-COND**

$*Place]_\sigma$

As was pointed out in 1.3, complex onsets and codas are also universally dispreferred, i.e. typologically marked. In fact, if a language allows complex onsets, it also allows simple onsets, if a language allows complex codas, it also allows simple codas. Hence, a constraint *COMPLEX has been postulated (PS:87). Since no intrinsic typological relation can be established between the presence of complex onsets and that of complex codas within one language, such a constraint has to be split into the two constraints in (42).[73]

(42) (a) **$*COMPLEX^{ONS}$**

$*[_\sigma CC$ ('Onsets are simple')

(b) **$*COMPLEX^{COD}$**

$*CC]_\sigma$ ('Codas are simple')

(from Kager 1999:97)

Many languages exhibit an absolute ban against complex onsets and/ or codas (i.e. they exhibit undominated $*COMPLEX^{ONS}$ and/ or $*COMPLEX^{COD}$), which may be implemented by repair strategies such as vowel epenthesis (see ch. 5) or consonant deletion.

Other languages allow complex onsets/ codas, provided the sequence of the consonants decreases in sonority from the nucleus towards the margins. To encode such sonority restrictions, PS (129) use two chains of 'anti-association' constraints, one banning specific segments from marginal positions (M), the other from peak positions (P). These are *M/a » *M/i » … » *M/d » *M/t and *P/t » *P/d » … » *P/i »

[73] According to whether the cluster to be avoided is in the margin (onset or coda) or in the nucleus, *COMPLEX may also be split into $*COMPLEX^{MARG}$ and $*COMPLEX^{NUC}$.

*P/a. Such chains can be subsumed under the cover constraint SONSEQ (or SSP, Sonority Sequencing Principle). A related constraint is HNUC, stating that a high sonority nucleus is more harmonic than one of lower sonority (PS: 16).

The inventory of syllable types allowed by individual languages is given by the interaction of such syllabic well-formedness constraints and segmental faithfulness constraints. That is, higher-ranked faithfulness constraints may prevent markedness constraints from imposing the preferred syllable structure CV, thus yielding more marked syllable types.

Cross-linguistically, some processes seem to apply in order to favour bimoraic syllables. In OT terms, they may be characterized by means of the high ranking of a constraint SYL-BIN. However, this constraint is often broken down into *3μ, militating against trimoraic syllables, and *σ_{μ},banning monomoraic syllables.

Al-Ageli (1995:151), however, points out that this latter constraint is rather unnatural, since it goes against the unmarked syllable type CV. Therefore, he proposes the constraint *Minimality Correspondence* ($\sigma \equiv$ FT), which requires that syllables be minimally bimoraic by transitivity, since they have to coincide with feet, which are themselves minimally bimoraic (see FT-BIN below).

As was mentioned in 1.3, whether a CVC syllable is light or heavy varies from language to language. In an OT perspective, it depends on the ranking of the constraint WEIGHT-BY-POSITION, stating that coda consonants are moraic (see ch.3). In languages in which CVC syllables are monomoraic, this constraint must be dominated by some other constraint, such as *2μ ('no bimoraic syllables') – which, considering what was said above, is rather counterintuitive – or *C_{μ} ('consonants must not be assigned moraic structure').

1.5.2.1.1 Italian syllable structure in OT

As was specified in 1.3, Italian allows for empty syllable onsets; therefore, ONSET must be dominated by faithfulness constraints such as DEP-IO (militating against epenthesis). *P/C, on the other hand, must be undominated, since consonants can never be syllabic. Since two intervocalic consonants are both syllabified into the following onset so long as this does not violate sonority sequencing, the ranking SONSEQ » *CODA » *COMPLEX^{ONS} must be postulated. *COMPLEX^{COD} is instead undominated (though complex codas do appear in loan-words), while *COMPLEX^{NUC} is ranked low. CODA-COND (*Place$]_\sigma$) is ranked rather high, though liquids and [s] may also appear in a coda. The cover constraint *OBS(place)$]_\sigma$ may thus be seen as undominated, with the exception of * s$]_\sigma$. *3μ also seems to be undominated.

1.5.2.1.2 Arabic syllable structure in OT

In most Arabic dialects, ONSET must be undominated, since onsets are always filled. In many varieties, including AA, *P/C is also unviolated. Moreover, in Levantine Arabic, as in Classical Arabic, both complex onsets and complex codas are disallowed; this is indicative of a high ranking of the two *COMPLEX constraints.[74] However, simple codas are permitted, and any consonant may appear in such a position. This means that *CODA is ranked low, and precisely lower than DEP-IO (or FILL) and MAX-IO (or PARSE-SEG), since neither vowel epenthesis or consonant deletion apply to repair a simple coda (see 4.3. and 5.3). *COMPLEX^{NUC} is also ranked low. SONSEQ is undominated in most dialects, though in some North African varieties it can be violated. *3μ is ranked high in AA: if final consonants are seen as extrametrical and PW-internal CVVC sequences as due to Adjunction to Mora (see 1.3), it may even be undominated.[75] *σ_μ, though not undominated, is also ranked high.

1.5.2.2 The foot and stress assignment

1.5.2.2.1 Footing

In the formulation of its constraints, OT still makes use of the same prosodic categories that were used in DT. Contrary to NV's model, however, in OT-based works the mora is often included, and the CG excluded, from the hierarchy, and binarity is heavily enforced.

The preference for all syllables to be parsed by feet is encoded in the constraint PARSE-SYL ('Syllables are parsed by feet').[76] More generally, constraints of the PARSE

---

[74] As was mentioned in 1.3, two-consonant clusters do appear word-finally, but the last consonant can be considered extrametrical.

[75] A theoretical framework like OT, however, can happily do without an ad hoc device such as Adjunction to Mora, precisely because *3μ , like all other constraints, is violable.

[76] Kager (1999:189) points out that PARSE-SYL has a dual interpretation:

i. If X forms a syllable, then X must be parsed by a foot (bottom-up)

ii. If X is not parsed by a foot, then X must not form a syllable (top-down) family require that prosodic constituents be parsed into the superordinate prosodic unit (e.g. PARSE-FT). Since such constraints are violable, it is clear that OT does not accept strict layering (see 1.2).

The anti-degenerate-foot constraint FT-BIN ('Feet are binary under moraic or syllabic analysis') states that the unmarked foot is formed either by two syllables or by one bimoraic syllable.[77] High ranking of PARSE-SYL and FT-BIN produces a complete parsing of syllables into binary feet, thus enforcing binary alternation. In unbounded foot systems, where feet comprise an unrestricted number of syllables, FT-BIN is ranked low. On the other hand, in languages in which FT-BIN » PARSE-SYL, a monomoraic syllable may remain unparsed. Such a syllable lies either at the left or at the right edge of the word, and its position is determined by another constraint, ALL-FT-L/R, by which 'every foot stands at the left/right edge of the PW'. In McCarthy and Prince (1993:94), these constraints are formulated as ALIGNFT-L/R ('Every foot stands in initial/final position in the PW'). When a word has more than one foot, they are clearly violated, since only one foot can stand at the edge of the PW (i.e., languages in which ALL-FT-L/R is undominated have only one foot per PW). They are gradient constraints, i.e. they take into account the summative distance of all feet from the relevant edge, measured in syllables. In English, for instance, ALL-FT-R (or ALIGNFT-R) must be ranked above both PARSE-SYL and ALL-FT-L (or ALIGNFT-L), as the tableau for the footing of *constriction* shows.[78]

(43)

Input: /kənstrɪkʃən/	ALL-FT-R	ALL-FT-L	PARSE-SYL
a. [kənstrɪk]ʃən	*!		*
b. [kəns] [trɪkʃən]	*!	*	
c. ☞ kəns[trɪkʃən]		*	*

[77] For those who allow for ternary feet, this constraint could be weakened to require that feet be binary or ternary.

[78] It may be objected, however, that the syllable /kəns/ is unparsed not because of ALL-FT-R, but because of a syllabic (rather than moraic) interpretation of FT-BIN. In fact, a four-syllable word like *breathalyzer* contains two feet, thus violating ALL-FT-R.

These constraints belong to the ALIGNMENT family, which comprises constraints that require that the edge of some constituent coincide with the edge of some other constituent.

Clearly, the ranking of ALL-FT-L with respect to ALL-FT-R determines the directionality of foot parsing. For instance, a language exhibits binary rhythm starting from the left edge of the PW if its constraint ranking includes FT-BIN » PARSE-SYL » ALL-FT-L » ALL-FT-R. In this case, the final, rather than the initial, syllable may remain unfooted.

Another constraint by which a PW-final syllable may remain unfooted is NONFINALITY ('No foot is final in a PW'), when NONFINALITY » PARSE-SYL. This constraint is the OT counterpart of DT extrametricality. Extrametricality of a word-final consonant, on the other hand, is given by high ranking of the constraint *FINAL-C-μ, by which the final consonant is weightless.

The 'functional antagonists' of ALL-FT-L/R are ALIGN-WD-L, by which 'every PW begins with a foot' and ALIGN-WD-R, by which 'every PW ends with a foot'.

1.5.2.2.2 Stress

Stress patterns, being characterized by conflicting forces such as *rhythm*, *culminativity* (by which constituents have a single prosodic peak), *quantity-sensitivity* and *edge-marking* (Kager 1999:142), are readily described in terms of OT constraints. Since the foot is the minimal unit of stress assignment, such constraints are often to do with foot parsing, so that most of them have already been mentioned above. Others, however, are linked to syllable structure. One such constraint is the WSP ('Weight-to-Stress-Principle'), by which 'heavy syllables are prominent in foot structure and on the grid' (PS: 53). Violation of WSP may lead to absence of secondary stress on heavy syllables. The mirror-image constraint SWP ('Stress-to-Weight-Principle') requires that stressed syllables be heavy. A related constraint is PK-PROM, by which the heaviest syllable within a domain is stressed.[79]

[79] Prince (1990) points out that 'these two halves of the purported biconditional "heavy iff stressed" [i.e. WSP and PK-PROM] have very different status in stress systems'. While PK-PROM's 'if stressed then heavy' and 'if light then unstressed' imply that heavy and stressed ≻ light and stressed, or that *light and stressed, WSP's 'if heavy then stressed' and 'if unstressed then light' are free from any such implication.

FT-UNIF(ORM) requires that syllables be parsed into feet in a uniform way with regard to weight. When this constraint dominates WSP, uniform weight-insensitive feet are formed.

A weaker version of the above-mentioned NONFINALITY states that no prosodic head is final in a PW, which may mean either that PW-final feet must not be the head of a PW or, more specifically, that PW-final syllables must be unstressed (PS: 40, 43, 52, 57).

The tendency for stress to fall at the edge of domains (demarcative property) is enforced by the two alignment constraints LEFTMOST ('The head foot is leftmost in the PW') and RIGHTMOST ('The head foot is rightmost in the PW').[80] Bidirectional stress systems are given by the interaction of LEFTMOST/RIGHTMOST with ALL-FT-L/R, and more precisely by the domination of peak alignment over foot alignment at the opposite edge (Kager 1999:168).

To summarize so far, a factorial typology of rhythmic and demarcative constraints will thus be as follows (from Kager 1999:171):

1. Non-iterative binary systems

   ALL-FT-X » PARSE-SYL = single-sided non-iterative

   ALIGN-WD-Y » ALL-FT-X » PARSE-SYL = double-sided non-iterative

2. Iterative binary systems

   FT-BIN, X-MOST » PARSE-SYL » ALL-FT-X = unidirectional

   FT-BIN, X-MOST » PARSE-SYL » ALL-FT-Y = bi-directional (simple)

   FT-BIN, X-MOST » PARSE-SYL » ALIGN-WD-Y » ALL-FT-X = bi-directional (complex)

3. Unbounded systems:

   ALL-FT-L, ALL-FT-R, PARSE-SYL » FTBIN

Within bounded systems, the rhythmic type of feet, i.e. their being right-headed (iambic) or left-headed (trochaic), is given by the following constraints:

(44) (a) **RHTYPE=I**

Feet have final prominence

[80] In PS, these constraints are part of the head-alignment constraints ALIGNHD, requiring that the leftmost/rightmost foot bear the main stress, i.e. that it is the head of the PW.

(b) **RHTYPE=T**

Feet have initial prominence

(from PS: 53)

As will be shown in more detail in section 3.1, the unmarked trochee is even, i.e. made up of two light syllables, while the unmarked iamb is uneven, i.e. formed by a light plus a heavy syllable. The two constraints EVEN-TROCHEE and UNEVEN-IAMB have thus been formulated.

The culminative property of stress, by which each prosodic constituent must have a single peak, is enforced by the combination of the requirement that each constituent in the prosodic hierarchy have one single DTE (DTE = 1) with the constraint GW=PW ('a Grammatical Word must be a Prosodic Word'). These constraints also determine the unmarked length of lexical items. This is two syllables, which is the most frequent size cross-linguistically, and the minimum size in many languages. Since, in other languages, the word minimum is two moras, the minimal Grammatical Word (GW) results by transitivity from the interaction between the two constraints GW=PW and FT-BIN and the requirement that each PW should contain at least one foot as its head (again, DTE = 1).

Finally, the constraint *CLASH enforces rhythmic alternation, by requiring that 'no stressed syllables are adjacent'.

1.5.2.2.3 Foot type and stress assignment in Italian

As was mentioned in 1.4, Italian feet are bounded and trochaic. This means that FT-BIN is ranked high, and that RHTYPE=T » RHTYPE=I.[81] Even if ternary feet are accepted, and thus the final syllable of a word like *súbito* 'immediately' is seen as parsed into a foot together with the two preceding syllables, PARSE-SYL must still be dominated by FT-BIN because unaccented initial syllables remain unparsed (e.g. *pa[gáno]* 'heathen'). The fact that words like *[còndi][ménto]* 'dressing' contain two feet shows that PARSE-SYL » ALL-FT-R, though ALL-FT-R still dominates ALL-FT-L (cf. *at[tènta][ménte]* 'carefully', and not **[àtten][támen]te*)[82].

[81] EVEN-TROCHEE, however, must be ranked low in Italian, because feet are often uneven.
[82] The same result, however, can be obtained by ranking ALIGN-WD-R high. Note that this constraint is undominated if ternary feet are allowed.

Since, in Italian, PW-stress always falls on the last foot, RIGHTMOST must be undominated. SWP is also ranked high, given that, apart from PW-final position, all stressed syllables are heavy.

1.5.2.2.4 Foot type and stress assignment in Arabic

Most varieties of Arabic, among which AA, exhibit a bounded weight-sensitive system, in which RHTYPE=T prevails. Stress always falls on the last foot in a word, and therefore RIGHTMOST must be undominated. Since prominence is assigned to heavy syllables whenever possible, WSP is also highly ranked. As was mentioned in 1.4, word-final syllables seem to be extrametrical, and therefore NONFINALITY must be ranked high, perhaps undominated. As far as footing is concerned, FT-BIN » PARSE-SYL, because monomoraic syllables may remain unparsed, as in *mu[Hámmad]* 'Mohammed'. Just as in Italian, however, PARSE-SYL » ALL-FT-R » ALL-FT-L (cf. *mu[màththi][líin]* 'representatives'), at least in varieties such as AA.[83] Again, if ternary feet are accepted (e.g. *[mádrasa]* 'school' rather than *[mádra]sa*), ALIGN-WD-R is undominated.

1.5.3 Morphophonology

Parallelism does not only mean that markedness and faithfulness constraints interact in a single hierarchy, but also that constraints pertaining to different linguistic modules, such as morphology and phonology, are evaluated in parallel. Therefore, OT rejects the assumption of traditional generative phonology that phonological processes apply after all morphological operations have applied. This assumption had already been challenged (although only partially) by Lexical Phonology (Kiparsky 1982a,b, Mohanan 1982, 1986), which showed that phonological processes may be morphologically conditioned. For instance, it has been pointed out (Mascaró 1976, Kiparsky 1982a) that root-internal elements are more resistant to change than elements in other positions, e.g. in an affix. This implies a distinction between *faithfulness* at large and *root-faithfulness*: the ranking *root-faithfulness* » *markedness* » *faithfulness* results in a situation in which a process is blocked within the root but

[83] For the problem of alignment constraints and directionality in the Arabic dialects, see ch. 4 and 5.

applies elsewhere (McCarthy and Prince 1995).[84] Another common assumption within OT is that morphemes are not linearly ordered in the input, and therefore LINEARITY-IO does not hold at morpheme boundaries (i.e. in derived environments).[85]

Within OT, another way in which morphology may influence phonology is by requiring that morphological constituents be maximally respected, both as regards the preservation of the contiguity of their constituent segments and the demarcation of their edges. The latter goal can be achieved by means of constraints that require the edge of a morpheme to coincide with the edge of some prosodic constituent, so as to signal it more clearly and consequently to facilitate perceptive processing.[86] Such constraints are part of the above-mentioned ALIGNMENT family, precisely because they call for grammatical constituents to be aligned with prosodic constituents.
The general format of Alignment constraints is as follows:

(45) **Generalized Alignment**

Align (Cat_1, $Edge_1$, Cat_2, $Edge_2$) $=_{def}$

$\forall$ Cat_1 $\exists$ Cat_2 such that $Edge_1$ of Cat_1 and $Edge_2$ of Cat_2 coincide.[87]

Where $Cat_1, Cat_2 \in$ ProsCat $\cup$ GramCat[88]

$Edge_1, Edge_2 \in$ {Right, Left}

(McCarthy and Prince 1993)

Thus, the above-mentioned ALL-FT-L/R can be formulated as ALIGN (Ft, L/R, PW, L/R), while RIGHT-/LEFTMOST can be formulated as ALIGN (Hd-Ft, L/R, PW, L/R).

[84] In spite of these considerations, it would be better to avoid the use of root-specific constraints, on the assumption that faithfulness constraints are blind to morphological structure.

[85] Hence ROOTLIN-IO, a linearity input-output constraint that refers only to non-derived environments.

[86] Other constraints, such as the above-mentioned GW=PW, establish a correspondence between a morphological category and a prosodic category. More generally, PS (43) propose a constraint LX $\approx$ PR *(Mcat)*, by which 'a member of the morphological category *Mcat* corresponds to a PW'.

[87] Note that the order in which the two categories appear is not random: Cat_1 is introduced by a universal quantifier, Cat_2 by an existential quantifier. The algorithm reads: 'for every Cat_1 there must be some Cat_2 such that the right/ left edge of Cat_1 coincides with the right/ left edge of Cat_2'.

[88] In individual languages, the grammatical category may also be filled by specific morphemes. In this regard, Kager (1999:119) points out that 'ALIGNMENT constraints may assume the function of designating an affix as either a prefix or a suffix, depending on the edge of the word (left/ right) with which it aligns.' This clearly weakens OT's assumption of universality, although the general format of ALIGNMENT constraints remains universal.

Although Alignment constraints have been proposed that align pairs of opposite edges, it is often assumed that they generally match pairs of identical edges (either both left, or both right). They are evaluated gradiently, measuring the distance between the edges referred to in the constraint. The unit used to measure such a distance varies from constraint to constraint, though the segment is by far the most common.

A different type of constraints concerning the edges of constituents belongs to the ANCHORING subfamily (L-ANCHOR, R-ANCHOR). Rather than requiring coincidence between the edges of different constituents, these constraints require correspondence between the segments at the edge of one constituent in the input and in the output. They thus prevent epenthesis and deletion from applying at the edge (see ch. 4-5).[89] Alignment and anchoring constraints may also replace the derivational concept of domain of application of a process: by requiring that the edge of a constituent coincide with that of another constituent or that it correspond with its own input, they prevent certain phenomena from taking place outside the boundaries of that constituent. In Tamil, for instance, geminate and sonorant consonants may appear in coda position, but only phrase-internally: phrase-finally, vowel epenthesis takes place after a geminate (e.g. /ʃiripp/ > [ʃirɪppɯ] 'laughter'), while coda sonorants are deleted (e.g. /sambandam/ > [sʌmbəndõ] 'relationship'). Wiltshire (1998:438-9) points out that this can be accounted for by high ranking of the constraint ALIGN-R(Phr,V), requiring that phrases end in a [+son, -cons] segment. Also highly ranked, but dominated by ALIGN-R(Phr,V), is ALIGN-R(Phr, Mor) ('Phrases end at the right end of a morpheme'), which prevents syllabification from applying across phrase boundaries and epenthesis from taking place at the end of phrases, except when it is imposed by the higher ranking constraint (ibid.:439).

1.5.4 Output-to-output correspondence

Correspondence constraints do not only express the relation between input and output, but may require identity between features in two output forms; that is, a surface form

[89] They are in fact correspondence constraints, and actually substitute for alignment constraints in correspondence theory.

is required to be identical in some respect to another free-standing surface form, called the *base*. This is termed Output-to-output correspondence (OO-correspondence), and makes use of the same types of constraints as IO-correspondence (LINEARITY, MAXIMALITY, DEPENDENCE, etc.). Typical examples of OO-correspondence are reduplication, where the reduplicant is required to be identical with (part of) its base, and morphological truncation, in which a truncated form copies phonological features off the corresponding non-truncated surface form.

When an OO-constraint dominates markedness and IO-constraints, underapplication or overapplication of a process may take place. For instance, truncated names in American English are not affected by the same regularities that other words with similar environments undergo: the truncated form of *Larry* [læri] is [lær] and not *[lɑr], that of *Sarah* [særə] is [sær] and not [sɑr], although the low unrounded vowel is always realized as [ɑ] before coda [r] in non-derived words (cf. *car* [kɑr], and not *[kær]). This behaviour can be attributed to high ranking of the OO-constraint IDENT-BT(back), requiring identity in backness between corresponding segments in the base and in the truncated form (for further discussion, see Kager 1999:259-65).

CHAPTER 2: GLIDE ~ VOWEL ALTERNATIONS

2.1 The issue

This section analyzes the issue of syllabicity alternation in vocoids, which, as was anticipated in chapter 1, lies at the heart of the definition of such a class. In the first subsection, an overview of the various types of alternations is given, starting with a more detailed account of the distribution of vocoids within the syllable, then going on to the question of what segments should be classified as glides, and finally dealing with Glide Formation and other types of glide alternation. Subsection 2.1.2 illustrates the related problem of the phonological status of glides, showing that there is indeed substantial evidence for a distinction between alternating and non-alternating vocoids, and introducing a few competing analyses of the issue. The final subsection offers an OT alternative to the above account.

2.1.1 Glide ~ vowel alternations

2.1.1.1 Syllabicity

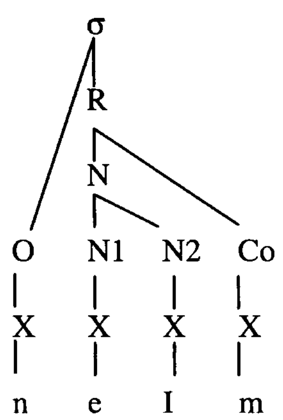
In section 1.1 a *glide* was defined as a non-syllabic vocoid, on the basis of the phonetic and phonological similarities existing between glides and vowels. Even though at the phonetic level glides have been shown to share characteristics both with vowels and with consonants, and their distribution within the syllable can to a certain extent be compared to that of the latter sounds, there are good reasons to believe that, at the segmental level, a glide like [j] actually exhibits the same feature matrix as the corresponding vowel [i]. The two classes of sounds can be differentiated only on the basis of their syllabicity.

In reality, vocoids can occupy more than two positions within the syllable, which led to the traditional distinction between vowels, on-glides and off-glides (cf. 1.3). As was pointed out in sections 1.1 and 1.3, off-glides, often transcribed as [i̯] or [u̯], exhibit very strong vocalic features, such as a high degree of openness and sonority, by which they are often analyzed as occupying a second position in the syllable nucleus rather than one in the syllable coda (O'Connor 1971, Lass 1976, Gimson 1989, Mioni 1993, Harris 1994). Moreover, long vowels can be represented as a vowel-(off-)glide sequence (see for example Donegan 1985:100-2), so that a long vowel like [i:] can be read as [ii̯] on the analogy with a sequence like [ai̯]. For these reasons, off-glides have even been classified as vowels, rather than as glides. This is confirmation that the line between glides and vowels is blurred, and that the higher or lower consonantality of a vocoid is merely due to its placement in the syllable.

If off-glides are placed in the nucleus, however, one should find a way to distinguish between long vowels (e.g. [ii̯]) and vowel-(homorganic) glide sequences (e.g. [ij]), even when the glide is only a surface realization. The latter configuration is easily explained if one sees the glide ([j]) as placed in the coda, rather than in the second part of the nucleus, of the syllable. We can thus identify four, rather than three positions within the syllable which can be occupied by a vocoid: the onset (O), a peak nuclear position (N1), a secondary nuclear position (N2) and the coda (Co). As a consequence, vocoids can be classified as O-vocoids (i.e. on-gides), N1-vocoids (i.e. syllabic vowels), N2-vocoids (i.e. nuclear off-glides) and Co-vocoids (i.e. coda off-glides).

Following these considerations, two positions within the syllable nucleus should be distinguished configurationally, in order to show that the first element in the cluster is more prominent than the second (see 1.3). A word like English *name* [neɪ̯m] will thus be represented as in (1).

(1)



In the diagram in (1), both N2 and N1 are dominated by N, i.e. are part of the nucleus. N1, however, has been linked to N by a perpendicular line, as opposed to the oblique line of N2, in order to show that the former is the more prominent – and indeed the only obligatory - node in the nucleus. Besides, the labels N1 and N2 have been used, rather than N' and N respectively, because N1 is not a higher projection of N2; rather, both N1 and N2 are projected from N, following a top-down approach.

However, this is not the end of the story. It has been pointed out that on-glides, too, may be syllabified in the nucleus rather than in the onset. Harris (1983) and Harris and Kaisse (1999), for instance, show that Spanish prepeak glides must, at least in some cases, be nuclear. More specifically, Harris and Kaisse (ibid.:126-130) argue that such glides are parsed as onsets when they are not preceded by tautosyllabic lower sonority segments, as nuclear elements otherwise. That is, high vocoids 'are onsets [only] if nothing better [i.e. no consonantal sound] is available' (ibid.:126).[1] To substantiate their claim, Harris and Kaisse (ibid.) use several diagnostics, which are briefly listed below.

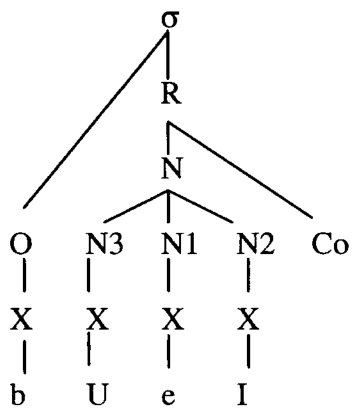
a. Consonantalization, i.e. a process that turns the palatal glide into an obstruent, only applies to syllable-initial /j/s, which can thus be argued to be more consonant-like O-glides.
b. *[CGVOs] syllables appear to be banned, while [GVOs] syllables are not. This cannot be explained by reference to segmental count within the syllable, because 5-segment syllables do exist in Spanish. It can be explained instead by assigning an upper limit of three segments to the rhyme. In the first instance the glide, being preceded by a tautosyllabic consonant, is syllabified in the nucleus, thus yielding an 'illegitimate' 4-segment rhyme; in the second case, on the other hand, the glide is syllable-initial and is thus syllabified in the onset, hence the perfectly acceptable 3-segment rhyme.
c. Post-consonant glides are often the result of diphthongization: since the underlying monophthongs are in the syllable nucleus, there is no reason to assume a successive movement of the glide into onset position, unless proven otherwise (see 2.2).
d. A process of j-deletion before the suffixes *–ito* and *–ista* takes place only when the glide is not syllable-initial. Since rhymes cannot contain sequences

[1] An analogous conclusion is reached by Rubach (1993b, 1998, in Harris and Kaisse 1999:130) as regards the distribution of prepeak glides in Slovak.

of identical segments, then the [j] must be nuclear only in this case (ibid.:127-8).

If at least some on-glides are syllabified in the nucleus rather than in the onset, then a further sub-syllabic node must be added to the diagram in (1). This is a prepeak nuclear position that we will conventionally call 'N3' (see (2)).

(2) Representation of Spanish *buéy* 'ox':



As was mentioned in chapter 1, only high vocoids can occupy all five positions. High vowels often alternate with non-syllabic glides, both synchronically (e.g. Spanish, Sanskrit, Ojibway) and diachronically (e.g. Classical Latin) (Donegan 1985:51). Non-high vocoids, on the other hand, are generally too sonorous to appear at the syllable margins, and often even as N2 elements.[2] Sonority is in fact the clearest factor affecting the capacity of a sound for syllabicity.

2.1.1.2 How many glides?

The number and quality of vowel-like sounds, or 'approximants', varies cross-linguistically, and many among them have been considered as glides by some scholars, while others have denied them such a status. English [ɹ], for example, can

[2] In languages like Italian (see 2.2), mid-high vowels such as [e] and [o] may optionally occupy an N2 position, as in *Paola* ['pao̯.la] ~ ['pa.o.la] 'Paula' or *aerazione* [ae̯.rat.'tsjo.ne] ~ [a.e.rat.'tsjo.ne] 'ventilation'. Ladefoged and Maddieson (1996:323) point out that a similar phenomenon applies – apparently more regularly than it does in Italian – in Nepali.

be considered as a glide on the basis of its articulatory features and especially because of its distribution (O'Connor 1971:74, Kahn 1980:120-3, Harris 1994:257-64). In non-rhotic varieties of English, [ɹ]-Insertion appears to constitute a systematic alternative to Glide Insertion (see below and chapter 5) in the case in which the preceding vowel is not high, and it has even been claimed (Ladefoged and Maddieson 1996:323) that, in rhotic varieties, it has a syllabic alternant [ɚ]. According to such authors (ibid.),

> [...] for many speakers of American English, the approximant ɹ at the beginning of the word *red* bears the same relationship to the vowel ɚ in *bird* as the approximant j in *yes* does to the vowel i in *heed.*

French [ɥ] (as in *huître* [ɥitʁ] 'oyster'), on the other hand, clearly constitutes an example of glide, paralleling [w] and [j] in its alternation with [y], even though this sound, like its homorganic vowel, is cross-linguistically marked. Even less common is the back unround glide [ɰ], found in Axininca Campa (Ladefoged and Maddieson 1996:322).

The glottals [h] and [ɦ] and pharyngeals [ħ] and [ʕ] have also been classified as glides, on the grounds of their low degree of oral stricture (cf. Chomsky and Halle 1968), but their status as such is more debatable (see 2.3).

In this thesis, I will therefore concentrate on more uncontroversial glides such as [j] and [w], i.e. on those glides that clearly alternate with homorganic vowels.[3]

2.1.1.3 Glide Formation

The fact that they often alternate with vowels is actually one of the strongest reasons to classify glides as vocoids. The phenomenon by which glides and vowels alternate in the same morpheme has traditionally been interpreted as a process of *Glide Formation* (henceforth GF), based on the fact that isolated or unaffixed forms generally display a vowel, which is thus seen as underlying (e.g. English *Italy* [ɪtəlɪ]

---

[3] As a matter of fact, authoritative phoneticians such as Ladefoged and Maddieson (1996:323-4) seem to consider 'semivowels' (they refuse to adopt the term 'glide') only those sounds which (may)

vs. *Italian* [ɪtæljən]).[4] According to this interpretation, such underlying high vowels lose their syllabicity – thus turning into glides – when followed by another vowel, and under certain other conditions (cf. 2.2 and 2.3). It is also easier to characterize this process as the formation of a glide from an underlying vowel and not vice-versa, since GF would apply in a more restricted environment – namely pre-vocalically – than *Vowel Formation* (VF) from an underlying glide. We will see, however, that in some cases a distinction between GF and VF has been claimed by some to have greater explanatory power (cf. 2.3).

Vowel ~ glide alternations are regulated by general principles of clash avoidance and (re)syllabification. By a widespread constraint, hiatuses - i.e. sequences of adjacent syllable nuclei - are typologically marked and therefore avoided. Such a constraint, which is evidently akin to the Onset Principle (see 1.3),[5] can be taken to be universal, though its implementation varies from language to language. In fact, in about half the languages of the world, hiatus is not allowed (Marotta 1987:859); in other languages, the ban is not so absolute, and sequences of vowels are simply dispreferred. The breaking of sequences of adjacent nuclei is thus a natural phenomenon, due to a more general Obligatory Contour Principle (OCP) effect enforcing symmetry by avoiding the clash between two analogous configurations. In order to break up the unwanted hiatuses, one of two things may happen: either (a) one of the two adjacent nuclei is deleted (as in elision, see ch. 4), or (b) the intervening empty onset is filled in by epenthetic material (as in Consonant Insertion, see 5.1). GF works both ways: by reducing the vocalicity of one of the two clashing sounds (the high vocoid), it allows it to occupy the empty onset, and at the same time it frees the corresponding nucleus that, being now empty, is deleted together with the whole syllable (cf. (3)).[6]

---

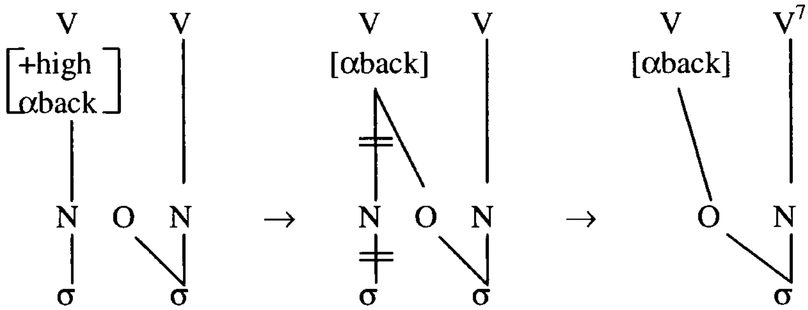
alternate with vowels, including English and Danish [r]. On this basis, the sub-class of semivowels can be distinguished from the more general class of approximants, of which it is part.

[4] In an approach that postulates a parametric location of initial syllabification (see 1.3), whether one can talk of GF or not will clearly depend on whether initial syllabification has already applied at that particular stage of the derivation in that particular language.

[5] Although, as was mentioned above, an on-glide may be syllabified in the nucleus rather than in the onset of a syllable, this latter possibility seems to be more viable in the case of nucleus clash avoidance.

[6] If resyllabification applies to the entire string, however, one cannot really talk of 'nucleus clash resolution': only one nucleus is supplied by resyllabification.

(3)



GF is thus a reduction process, and often applies postlexically. Ultimately, processes of this kind can be traced back to ease of articulation requirements, which are a strong trigger to connected speech phenomena.

When talking of alternations such as these, it must be borne in mind that, according to this approach, homorganic vocoids share the same set of distinctive features, and differ only with respect to the position they occupy in the syllable. Therefore, they do not alternate at the segmental, but at the syllabic level.

Since an off-glide is placed in the syllable nucleus, its presence in an intervocalic context does not in itself constitute hiatus-breaking, but actually triggers the formation of an O-glide, just as the presence of a full vowel does. It should be noted, though, that GF from an off-glide does not delete the first clashing nucleus, which is still occupied by the preceding vowel, but separates the two nuclei by means of a non-nuclear element.

GF does not apply every time a clash takes place between a high vowel and another vowel: its application is subordinate to the setting of certain parameters. One of the factors that can block GF is stress placement. A stressed high vowel, being longer and more sonorous, often remains syllabic even when followed or preceded by another vowel. Lack of prominence on the affected segment is in fact a typical requirement for all phenomena of vowel reduction (also see chap. 4).

Although GF is a widespread process, its domain of application varies considerably from one language to the other. In some languages, such as English and Dutch, the operation is limited to apply within the PW, and is blocked outside of it. In other languages, such as Spanish and Italian (see 2.2), it applies in larger domains and is a connected speech phenomenon.

---

[7] 'V' here stands for 'vocoid', i.e. it is short for [-consonantal].

2.1.1.4 Other processes of glide derivation

Another operation by which glides can be derived – and another way of resolving a vowel-clash - is Glide Insertion, by which an O-glide is inserted between two adjacent vowels. Glide Insertion entails no alternation with a vowel, hence the process will be dealt with later on in the thesis (chap. 5).

Glides can also be derived by means of diphthongisation, which is the alternation between a (not necessarily high) vowel and a diphthong. It is a process of fortition which appears to take place mainly in the lexicon (Di Luzio 1988:308), and which may affect either a long or a short vowel, re-enforcing it by the addition of a glide, or by transforming it into a vowel-glide sequence (cf. Andersen 1972). Diphthongs can in fact be defined as a sequence of two vocoids belonging to the same syllable. They are traditionally classified into *rising* – composed by an on-glide plus a vowel, as in [ja] – and *falling* diphthongs – composed by a vowel plus an off-glide, as in [ai̯]. According to some phonologists (e.g. O'Connor 1971:107-10), only these latter can be labelled as diphthongs proper, and 'rising diphthongs' are better viewed as CV sequences, since on-glides are generally placed in the syllable onset.[8] However, it seems to be appropriate to draw a distinction between CV and GV clusters, since vowels typically alternate with a vowel + off-glide or with an on-glide + vowel sequence, but not with a CV sequence.

Diphthongisation can either be diachronic or synchronic. In this latter case, an underlying vowel (usually a long one) surfaces as a diphthong in certain circumstances, generating a glide. Diachronic diphthongisation, on the other hand, applied at one point in the development of a given language, yielding diphthongs which are now fossilized and do not alternate with monophthongs any longer (e.g. Old Icelandic *[ska:l]* 'bowl' > Modern Icelandic *[skau̯l])*. Only synchronic diphthongisation is relevant here.

---

[8] Donegan (1985:190) points out that falling diphthongs, being exhaustively contained in the syllable rhyme, function as a unit for weight and rhyming purposes, while rising diphthongs, the first element of which is in the onset, do not. Rising diphthongs exhibiting nuclear glides, however, behave like falling diphthongs, since nuclear on-glides appear to be moraic (Harris and Kaisse 1999:130, following Rubach 1998).

2.1.2 The phonological status of glides

The existence of alternations leads to the fundamental issue of the phonological status of glides, i.e. to the question of whether all glides are derived allophones of the homorganic high vowels or whether, in some languages at least, there are underlying, phonemic glides. Such an issue has been at the centre of a debate which goes back to Jakobson, Fant and Halle (1952), and even to Grammont (1933), who maintained that there is no underlying distinction.

For many languages, it has been persuasively argued that all glides are allophones of the corresponding high vowels. In Romanian, for example, all glides appear to be derived, since their distribution is entirely predictable on the basis of such factors as the quality of adjacent segments, stress patterns, register, the internal organization of the lexicon, the degree of nativization of a word and its frequency of usage (Chitoran 1997). Similar considerations have been made for Malay (Durand 1987:99) and English (cf. O'Connor 1971:74-81; Jones 1972:206-11).[9]

2.1.2.1 Diagnostics

For other languages the existence of underlying glides has been claimed, and different diagnostics have been used as evidence. Adopting a Natural Generative Phonology perspective, by which the UR of a morpheme must be identical to one of its phonetic realizations (cf. Lyche 1979:319), would of course simplify things, because for the forms which always appear with a glide we would have to postulate one underlyingly. However, since in the present approach glides and homorganic vowels share the same set of distinctive features and alternations between them only take place at the syllabic level, it cannot be a suitable solution to theorize that a segment in a specific morpheme is underlyingly prelinked with a certain syllabic position only for the reason that it does not get a chance to surface in any other position. The vocoid /I/ in the word *yearn*, for instance, will always appear in prevocalic position – and therefore always surface as an O-glide – whatever affix or lexeme may follow or precede such a root.

[9] For further considerations on the phonological status of glides in English, see Ladefoged and Maddieson (1996:322-3), Bosisio (1998:28).

There are other, more convincing ways of testing the phonological status of glides in a given language. First and foremost, one should try to prove the validity of the opposition between a glide and the corresponding vowel in minimal pairs: no other factors, such as the distribution of stress or the presence of morphological boundaries, should be responsible for the distinction. Harris and Kaisse (1999:123-4), for instance, claim that Spanish exhibits clear (quasi-)minimal pairs, in which the contrast between vowel and homorganic glide in analogous environments is not attributable to morphological factors and is not a matter of free variation (e.g. *vac[i]aba* 's/he emptied' ~ *vic[j]aba* 's/he vitiated', *s[u]eco* 'Swedish' ~ *z[w]eco* 'wooden shoe').

GF is also a good indicator of the status of glides: in an environment in which this process is normally blocked, an underlying glide would obviously still surface as such (cf. Hannahs 1998a, b; Rea 1979:357-8). Moreover, in front of an underlying glide words may assume a different shape to the one they assume in front of a derived one (cf. liaison in French,[10] distribution of articles in Italian).

An excellent diagnostic of the phonological independence of glides is the insertion of a syllabic sound to break up a glide + consonant cluster: if such a glide were not underlyingly non-syllabic, it would turn into a vowel rather than requiring the insertion of one. This is precisely what happens in a language like Temiar. If the first glide of each of the forms of the root /kɔ:w/ 'call' in (4) were underlyingly syllabic or underspecified for syllabicity, it would surface as a vowel; instead, it calls for the insertion of the segment in bold.

(4)

/kwkɔ:w/ > [k**e**wkɔ:w] (not *[kukɔ:w]) active, continuative

/kwnɔ:w/ > [k**e**wnɔ:w] (not *[kunɔ:w]) active, continuative, nominalised

/trwkɔ:w/ > [tər**e**wkɔ:w] (not *[tərukɔ:w]) causative, continuative

(from Itô 1989:252-3)

---

10 According to Hannahs (1998a:5-6), French liaison applies before a derived glide (underlyingly a vowel) but is blocked before a non-derived one; thus, *des whiskys* [de wiski] 'some whiskies' would have an underlying glide, *les oiseaux* [lezwɑzo] 'the birds' a derived one.

2.1.2.2 The case of French

Not all languages, however, exhibit such clear evidence. It is often difficult to say whether there are underlying glides, because the overwhelming majority of semiconsonants are still in complementary distribution or in free variation with the high vowels, and the validity of the few minimal pairs is often moot. Such is the case of French, which has been studied and debated by phonologists over the last few decades. As was mentioned above, besides [j] and [w], this language has a third glide [ɥ], corresponding to the front round high vowel [y]. All such glides can be derived by means especially of GF; for some of them at least, however, it has been claimed that they can be underlyingly non-syllabic in specific cases. Spence (1971:207) points out that the selfsame fact that there exist high vowel + vowel sequences shows that glides are not the automatic realization of high vowels before another vowel. However, his claim can easily be invalidated if the maintenance of a hiatus is shown to be due to factors such as stress placement or the presence of a particular morphological boundary. The existence of minimal pairs, or quasi-minimal pairs, such as *paye* [pɛj] 'wages' and *pays* [pɛi] 'country', *pied* [pje] 'foot' and *piller* [pi(j)e] 'to loot', *voix* [vwa] 'voice' and *voua* [vua] '(s/he) devoted', *trois* [tRwa] 'three' and *troua* [tRua] '(s/he) perforated', *truite* [tRɥit] 'trout' and *truelle* [tRyɛl] 'trowel' (Hannahs 1995:1130, 1998a:3; Rea 1979:356; Spence 1971:204) can actually be questioned on a similar basis: *voua* and *troua* exhibit a morphological boundary which is absent from *voix* and *trois* and the presence of a [j] in *piller* is not due to a glide insertion, but it is the rendering of the grapheme <ll> (see below). In the word *pays* [pɛi], on the other hand, the final [i] is stressed, which could be the cause of the blocking of GF.[11] The presence of a pair like *truite/ truelle* is certainly more difficult to account for without recurring to the notion of underlying glide.[12]

In fact, Hannahs (1998a:24) points out that words like *truite* or *trois* must have an underlying glide, since a preceding obstruent + liquid (OL) cluster usually blocks GF. However, Spence (1971:206) points out that, in the sequences OL + /Y/ + [i] and OL + /U/ + [a] (or [ɛ̃]), the penultimate sound almost invariably surfaces as a glide,

[11] It can, however, be argued that it is precisely because [i] is a vowel that it is stressed, taking into consideration that in French main stress always falls on the last vowel in the word, and assuming that initial syllabification takes place before stress assignment (cf. Italian and Spanish in 2.2).

which seems to indicate that the whole issue may be more a matter of coarticulation than of phonemic distinction. Moreover, Rea (1979:359-60) shows that the diphthong [wa] (written ⟨oi⟩) often alternates with [e] (e.g. *loi* [lwa] 'law' / *légal* [legal] 'legal'), thus claiming that the UR may in fact be /e/ and the glide be derived by a process of diphthongisation similar to that which yields [j] or [w] in 'movable diphthongs' (cf. 2.2).[13] Moreover, some surface word-final [j]s may well have /l/ as their UR (e.g. *aller* [ale] 'to go' / *aille* [aj] '(I) go (subj.)'. The same could be said for words like *paille* [paj] 'straw' and *rouille* [Ruj] 'rust', although their Co-glides do not alternate with [l]. The derivation of such glides from consonants, however, though certainly real diachronically, can easily be questioned at the synchronic level. It must not be forgotten that, in time, URs can undergo restructuring, so that a diachronically derived glide can become underlying in fossilized forms (Rea 1979:361).

2.1.2.3 Consonantal and vocalic glides

In 1.1, however, the possibility was mentioned of tracing a distinction between consonantal and vocalic glides. According to such a distinction, semiconsonants like those in French *rouille* or *aille*, or like that in Spanish *paella* [paɛja], or like the [w] in the above-mentioned Temiar examples, could be construed as contoids, rather than as vocoids. In this sense, the assumption cannot be excluded that their diachronic history may have affected their present status. The issue of vocalic and consonantal glides, however, is a dangerous ground that demands precaution: the main risk is to adopt the distinction as an *ad hoc* solution to the problem of the phonological status of glides, by simply labelling non-alternating glides as [+consonantal]. Moreover, since [+consonantal] sounds can be syllabic in many languages, assigning such a feature to 'consonantal' glides does not necessarily explain why they never surface in syllabic position.

2.1.2.4 Berber and partial underlying syllabification

A better solution is to keep the same feature structure for alternating and non-alternating glides, but to have partial underlying syllabification (see Archangeli

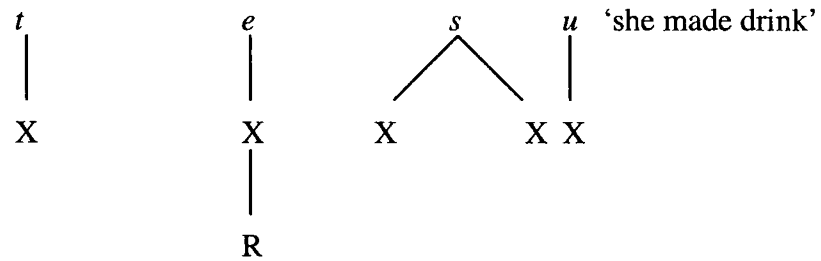
---

[12] *Truelle* does not exhibit a derived environment in French, unless it has preserved the original diminutive function of its Latin ancestor *trulla* (< *trua* 'stirring spoon, ladle'), which is highly unlikely.

1984), restricted to non-alternating segments. This is the analysis that Guerrsel (1986) proposes to account for the intriguing patterns of the Ait Seghrouchen dialect of Berber (henceforth, ASB). On the basis of most compelling data, the author shows that the elimination of the feature [syllabic] and the attribution of the same feature structure to homorganic glides and vowels does not invalidate the claim that, in some specific cases, a distinction must be drawn between glides and high vowels underlyingly.

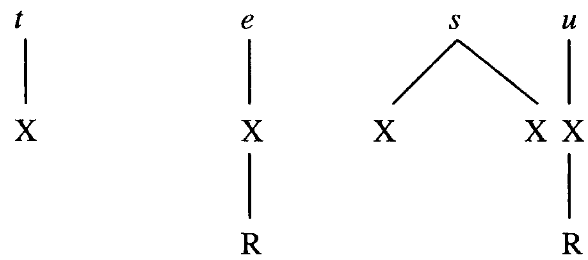
ASB has two glides: [j], as in *aryaz* [arjaz] 'man', and [w], as in *amawal* [amawal] 'dictionary'. These are in frequent alternation with the high vowels [i] and [u] respectively, as in *[i]-ru* 'he cried'/ *[j]-ari* 'he wrote', *[u]-mazan* 'messenger'/ *[w]-ansa* 'place'. However, high vowels do not always alternate with glides, as in *tsell-[i]* 'she heard'/ *tenna-[j]-[i]* 'she told me' (not **tenna-[i̯]*), with Glide Insertion rather than GF. Minimal pairs can also be found, such as *tess[w]-ax* 'she made us drink' (not **tess[u]-[j]-ax*)/ *tess[u]-[j]-ax* 'she made us a bed' (not **tess[w]-ax*). Clearly, there are two sets of high vowels, alternating and non-alternating, thus an underlying distinction must be drawn, as a function of syllable structure. Guerrsel concludes that

a. Alternating vocoids are underspecified for syllabicity at the phonemic level, and are assigned an onset (O) or rhyme head (R) node by position, presumably at the end of the lexical cycle:



b. Non-alternating vocoids are underlyingly pre-linked with a rhyme head (R):

[13] S.J. Hannahs (personal communication) points out, however, that not all such [wa] alternate with [e]



'she made a bed'

In other languages, of course, the contrast between alternating and non-alternating vocoids may not be as clear as in ASB, or may not exist at all. Nevertheless, Guerrsel's analysis leads to a universal consideration: since the syllabicity of a segment is a function of its placement after syllabification, and not one of its intrinsic features, in the unmarked case a vocoid is underlyingly underspecified for it. Non-alternating vocoids, therefore, should be somehow marked in their URs. This can be done by pre-linking them to a specific position in the syllable.[14]

An analogous position is assumed by Harris and Kaisse (1999:133) as regards non-alternating high vowels in Spanish. They argue that, in the unmarked situation, Spanish high vocoids are syllable peaks only when not adjacent to other more sonorous vowels, while hiatuses ought to be treated as special cases by pre-linking them to a syllable head N ('R' in Guerrsel 1986).

It has been claimed, however, that the fact that not all segments derive their syllabicity from syllabic position damages the theory. Noske (1993:94), for instance, points out that referring to an element as unspecified for syllabicity while maintaining a specification as syllabic and non-syllabic for other elements makes the theory too powerful, and is actually equivalent to introducing a third feature value. If, however, partial underlying syllabification is parameterized and limited to cases in which there is a three-way distinction between always syllabic, always non-syllabic and alternating segments, it becomes an effective way to characterize an actual reality of natural languages.

(e.g. *moi* ~ *m[ə]* 'me', *bois* ~ *b[y]* 'drink', *trois* ~ *tr[i]* ~ *tr[e]* 'three').

[14] In dealing with a similar case of alternating and non-alternating vocoids in Kikamba, Roberts-Kohno (1999) advances the suggestion that lack of alternation is due to the presence of an empty C position on the skeleton intervening between the clashing vowels, rather than to a difference in vocalicity or

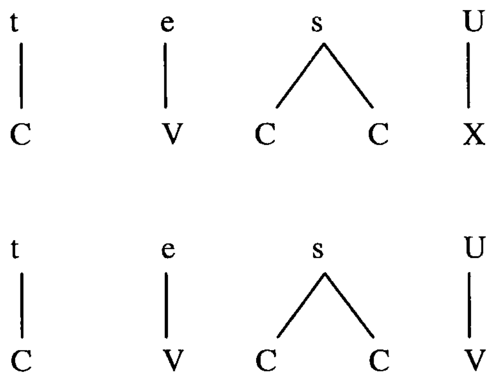
2.1.2.5 The CVX-skeleton

Although it easily accounts for the presence of segments that always surface in N1 position, partial underlying syllabification cannot characterize segments that always occupy marginal positions, as in the case of glides that do not alternate with vowels, as in French *rouille* [Ruj] 'rust' ~ *rouillé* [Ruje] 'rusty'. Notice that the glide here does alternate between a Co and an O position, so that the solution cannot lie in pre-linking it to a specific sub-syllabic node.

This problem can be obviated if the nature of segments with respect to their 'syllabificatory disposition' is specified on the skeleton. That is, segments which are obligatorily syllabified in marginal positions (O and Co) correspond to a C slot on the skeleton, segments that are obligatorily syllabified in the nucleus correspond to a V slot, while alternating segments are linked to an X slot (see 1.2, 1.3).

The above-mentioned ASB verbs would thus be represented as in (5).

(5)



(a) 'she made drink'

(b) 'she made a bed'

Another advantage of this approach is that, since the connection between the CVX-slots and the melody varies cross-linguistically (e.g. in one language /w/ or /n/ may be specified for C, in another they may alternate), the skeletal slots are not redundant. The CVX-skeleton, however, is no better than partial underlying syllabification in its introduction of a three-way distinction between C, V and X segments. Moreover, the use of partial underlying syllabification has the advantage of showing that non-alternating high vocoids are marked, while alternation is the unmarked case (Harris and Kaisse 1999:124).

---

syllabification levels. Establishing whether this postulation can be extended to ASB would require an in-depth study of other phenomena, which lies outside the scope of this thesis.

2.1.3 Glide distribution and Glide ~ Vowel alternation in OT

As was mentioned above, hiatuses are universally dispreferred, but only in certain languages are they completely banned. This generalization is easily captured in OT by means of a constraint that disallows hiatuses, called NoHiatus. Such a constraint actually performs an analogous function to Containment Theory FillOns or Correspondence Theory Onset, declaring that syllables must have onsets (see 1.5). Since, however, high vocoids may alternate between an N1 and an N3 position – rather than between an N1 and an O position – the more general NoHiatus will be used throughout this thesis.

On the other hand, the constraint *Complex militates against complex onsets and/ or codas, thus favouring VF when one of the clustering marginal sounds is a glide. SonSeq, requiring the sonority of a string of segments to increase as it approaches the syllable nucleus, also contributes to this effect, since it bans a highly sonorous vocoid from being surrounded by lower sonority consonants without being syllabic.

The constraint that prevents vocoids from associating to marginal positions in the syllable is *M/V (see 1.5). When this is outranked by NoHiatus, GF may apply. *M/V should actually be split into different constraints, such as *M/a, *M/o, *M/i, and so on. Constraints such as *M/a and *M/o are generally ranked higher than NoHiatus, Max-V-IO (militating against Vowel Deletion) and/ or Dep-C-IO (militating against Consonant Insertion), thus preventing GF from non-high vowels; constraints such as *M/i or *M/u, on the other hand, are often ranked lower than NoHiatus, Max-V-IO and Dep-C-IO, thus allowing GF from high vowels. If so-called 'underlying' glides are linked to a C, then *P/C (no contoids in peak position) is also relevant, because it militates against VF.

As was mentioned in 2.1.1.3, stress placement often blocks the application of GF. This is evidence for a constraint Pk-Max-IO, independent of *M/V.

(6) Pk-Max-IO:

Let α be a segment in the input and β its correspondent in the output. If α is the stress peak of the input form, then β is the stress peak of the output form.

(From Kager 2000:127)

When the high vowel is stressed, if such a constraint outranks NoHiatus, the hiatus is maintained. However, if both constraints outrank Dep-C-IO, a non-syllabic sound is inserted between the two clashing vowels.

In OT, the domain of application of GF, as that of many other phenomena, is determined by the ranking of alignment constraints. In languages such as Dutch or English, where GF does not apply outside the PW, the constraint Align R(PW, σ) is ranked high.

A constraint that could work against monophthongization is Uniformity-IO, by which 'no element of the output has multiple correspondents in the input ('no coalescence')' (McCarthy and Prince 1995). A similar uniformity constraint, stating that no element in the input has multiple correspondents in the output, should be postulated against diphthongisation. In order to distinguish it from the constraint against coalescence, this can be called *Fission. Moreover, since only about 1/3 of the world's languages have diphthongs at all (Ladefoged and Maddieson 1996:321), some markedness constraint must be militating against them. Presence or absence of falling diphthongs is in fact regulated by the ranking of $\text{*Complex}^{\text{NUC}}$. Presence or absence of rising diphthongs is regulated by the same constraint if on-glides are part of the nucleus; otherwise, it may be regulated by constraints on sonority distance. A cover constraint *Diphthong may also be used.

The problem of the UR and of the phonological status of glides should not exist in OT because of so-called 'Richness of the Base' (see 1.5). However, since specific vocoids may alternate in certain words and not in certain others, whether a segment always occupies a marginal or a nuclear position, or whether it may alternate must be specified in the lexicon, i.e. in the input, thus weakening the notion of Richness of the Base.

2.2 Alternations in Italian

This section offers an account of the distribution of glides and vowels in Italian. The first subsection introduces Italian glides and their differences in vocalicity. The second one analyzes the processes of glide ~ vowel alternation, dealing first with diphthongization and then with so-called Glide Formation (GF) and its relationship with stress assignment and (re)syllabification. Finally, the third subsection tackles the problem of the phonological status of Italian glides. Each issue is analyzed both in Derivational Theory (DT) and in OT terms.

2.2.1 Italian glides and their vocalicity

Standard Italian (SI) has only two glides: the palatal glide [j] and the labial-velar glide [w].[15] No other segments in the inventory of this language can be considered as part of this category. The two glides can be realized both as on-glides and as off-glides, but to my knowledge there are no examples of Co-glides.[16] The palatal on-glide appears in words such as *ch[j]esa* 'church', *opera[j]a* 'workwoman' or *[j]odio* 'iodine', the labial-velar on-glide in words such as *ling[w]a* 'tongue', *p[w]erile* 'childish' or *[w]omo* 'man'. The palatal off-glide is found in words such as *anda[i̯]* '(I) went', *no[i̯]* 'we' or *intu[i̯]to* 'intuition', the labial-velar off-glide in words such as *fa[u̯]no* 'faun' or *re[u̯]matisma* 'rheumatism'. Homorganic vowel-glide or glide-vowel sequences are disallowed in Italian, except when the second segment is in N2 position, yielding a long vocoid.

It has been claimed, however, that the two Italian glides exhibit a considerably different degree of vocalicity, which may reflect on the phonological level (Marotta 1987:878). The labial-velar vocoid, which is also less frequent and appears in more restricted environments,[17] seems to be more vowel-like than the palatal vocoid. This claim can be supported by means of different diagnostics, one of which is the distribution of the definite and indefinite articles. These take on different forms, according to the initial sound(s) of the following word. The two glides behave differently with respect to the shape of the articles that can precede them. While a word beginning with [w] takes the pre-vocalic forms *l'/ un / un'/ gli* (cf. *l'[w]ovo* 'the egg', *un [w]omo* 'a man'), one beginning with [j] takes the forms which appear before a specific set of consonants or consonant clusters - namely <s> + consonant,

---

[15] As was mentioned in 2.1, in Italian the high-mid vocoids /e/ and /o/ may optionally be non-syllabic; however, when non-syllabic, they can only be N2-glides.

[16] Co-glides can only be found in certain non-standard varieties, such as Neapolitan Italian (e.g. *Maria* [ma.rij.jə] 'Mary').

[17] Across word boundaries, the scarcity of [w] + V sequences is also due to the rarity of word-final /U/s in a language like Italian, where Latin /U/ has turned into /o/ in such a position. Moreover, when an /U/ appears at the end of a word, it is virtually always stressed, as in *Perù* 'Peru', *più* 'more' or the strong monosyllable *tu* 'you', and therefore cannot be realized as a glide (see below).

[ts], [dz], [ɲ], [ʃ], [ʎ], [ks], [pn], [pt], [ps], [kt], [mn] and [ft] - i.e. the forms *la/una/lo/uno/gli* (cf. *lo [jj]odio* 'the iodine', *uno [jj]ogurt* 'a yoghurt')[18].

Besides the articles, there are other words whose form reveals the difference in vocalicity between [j] and [w], such as the preposition *di* 'of', which is elided before a [w] but not before a [j], or the modifiers *bel/bello* 'nice (masc.sg.)' and *quel/quello* 'that (masc.sg.)', which behave exactly like the definite article (see sections 4.2 and 5.2).

An interesting case is that of the conjunction *e* 'and', which optionally takes the form *ed* with a linking /d/ when followed by a vowel-initial word. According to a native speaker's intuitions, a co-ordinate structure like *giardini ed orti* 'gardens and kitchen-gardens' is perfectly grammatical, while one like *dittonghi ed iati* 'diphthongs and hiatuses' sounds odd and clumsy. An expression like *donne ed uomini* 'women and men', on the other hand, stands somewhere in between: it seems to be quite acceptable, but would not generally be used by a native speaker. This appears to show that [w] is less sonorous than a full vowel but definitely more sonorous than the palatal glide.[19]

These considerations may induce an interpretation of the on-glide [w] as placed in the rhyme rather than in the onset (cf. Harris's (1983) analysis of Spanish on-glides, in 1.3). However, co-occurrence with other segments seems to point in the opposite direction. In Italian the glide [w], just like [j], must belong to the onset, because initial clusters like [stw-] (e.g. *stuolo* 'crowd, host') or [str-] (e.g. *strano* 'strange') are allowed, while clusters like [strw-] (or [sprj-]) are not. Moreover, Italian words like *cli.en.te* 'customer' or *tri.on.fo* 'triumph' are trisyllables, i.e. a high vocoid following an OL cluster generally[20] surfaces as a vowel, while Spanish *clien.te* and *triun.fo*, at least according to Harris (1983:9), comprise only two syllables. Finally,

[18] Notice that [j], like the other palatal consonants, is doubled in SI, i.e. the variety in which the different forms of the articles are phonologically motivated (see 4.2 and 5.2); contrary to the other doubling consonants, however, [j] is only geminated across word boundaries, and not within PWs (e.g. *sella[j]o* 'saddler', and not **sella[jj]o*).

[19] The validity of all these tests, though, has been challenged by several linguists. O'Rourke (1992), for example, suggests that the distribution of the article *lo* has no real phonological motivation, but is the result of the grammarians' attempts to impose some order on the paradigm. Spence (1971:200) also criticizes the idea of looking at the forms of the determiners as a diagnostic for the vocalicity of glides.

there are indeed w-initial words that take the form of the articles which normally precedes simple consonants, i.e. *la/una/il/un/i*. These are often loan-words, such as *la World Wide Web*, *il Word*, *i whisky*. Besides, when preceded by /s/, [w] tends to behave like a consonant in loan-words like *swatch*, in which the /s/ assimilates in voicing yielding [zwɔtʃ] and the preceding article takes on the form *lo/uno/gli* (see above); in native words like *suono* [swɔno] 'sound', on the contrary, the /s/ does not assimilate, and is preceded by the variants *il/un/i*, as if the [w] were more vowel-like.[21] This observation, though, could not be extended to [j]-initial words, since these are always preceded by pre-consonantal forms of the articles, independently of their origin. On the other hand, when preceded by /s/, the palatal on-glide behaves like a vowel, so that a word like *siero* 'serum' is realized as [sjɛro] rather than as *[zjɛro], in clear contradiction with what was said above.

However, a closer look at the two groups of words – those in which on-glides behave like consonants and those in which they behave like vowels – reveals that a factor other than their origin or the quality of their glides sets them apart. The glide in *uomo, uovo*, *suono* or *siero* is part of a so-called *dittongo mobile* 'movable diphthong' (see 2.2.2.1 below), i.e. of a diphthong which alternates with a vowel according to stress distribution (cf. *uómo* 'man' ~ *omíno* 'little man', *uóvo* 'egg' ~ *ovále* 'egg-shaped', *suóno* 'sound' ~ *sonóro* 'sonorous', *siéro* 'serum' ~ *serología* 'serology'). All other examples given above, i.e. those words in which on-glides behave like onset consonants, do not display movable diphthongs. It seems therefore legitimate to hypothesize that glides in movable diphthongs, which derive synchronically from nuclear vowels, may be placed in the nucleus, while all other Italian on-glides may be in the onset. The case of Italian thus seems to differ from that of Spanish, where on-glides are in the nucleus when preceded by tautosyllabic

[20] Just like in the case of French (see 2.1), there seem to be some exceptions to this pattern, such as *truogolo* [trwɔ.go.lo] 'trough'. In a DT perspective, these may be interpreted as displaying an underlying glide (see 2.2.3).

[21] In DT, rather than to a different degree of vocalicity, this divergence may perhaps be attributed to a distinction between underlying and derived glides: in loan-words, the [w] is already present in the UR, so that it behaves like a consonant, while in native words the labial-velar glide is derived either by means of diphthongisation or of GF, and is thus underlyingly a vowel. In OT, where this alternative is not available, a difference in vocalicity between the two occurrences of [w] must be postulated, so that a constraint like S-VOICE – by which an /s/ assimilates in voicing to the following consonant – only applies to loan-words.

consonants, in the onset when syllable-initial. A more similar case is that of French, in which high vocoids must be lexically distinguished as to whether they may fill onsets, since a glide-initial word may take either the pre-consonantal or the pre-vocalic form of the definite article (cf. *le whiskey* [lø wiski] 'the whisky' vs. *l'oiseau* [lwazo] 'the bird', *le yod* [lø jɔd] 'the yod' vs. *l'iode* [ljɔd] 'the iodine') (Tranel 1987:117, in Harris and Kaisse 1999:130; also see 2.1, fn.10). French, though, contrary to Italian, does not display movable diphthongs, and thus the division between onset and nuclear on-glides seems to be more arbitrary.

In spite of what has just been said, Italian high vocoids appear to exhibit some difference in vocalicity even when movable diphthongs are not involved. This is reflected in the way GF patterns: /U/ can sometimes be deleted in a prevocalic environment, while /I/ always seems to form a glide. A sentence like (7a) may be pronounced as in (7b), while a sentence like (8a) is normally pronounced as in (8b).

(7) (a) Radu Lupu è a Genova.
'Radu Lupu is in Genoa'
(b) Radu Lup' è a Genova.

(8) (a) Pollini è a Genova.
'Pollini is in Genoa'
(b) Pollin[j] è a Genova.

In OT terms, these considerations lead to the conclusion that *M/U (no /U/ in marginal positions) is ranked higher than *M/I (no /I/ in marginal positions). More precisely, since deletion rather than GF applies to /U/ but not to /I/, *M/U » MAX-V-IO » *M/I. This, however, is not always the case: GF does often apply to /U/, too, and, alternatively, the hiatus may even be maintained. That is, forms with [u], [w] and [0] are in free variation. As was mentioned in 1.5, perhaps the best solution to the problem of free variation in OT is resorting to free ranking. In this case, this means that the three constraints *M/U, MAX-V-IO and NOHIATUS are ranked in different ways with respect to each other in each of three subhierarchies: in the first subhierarchy, *M/U is ranked lowest, thus producing GF, in the second one, MAX-V-IO is ranked lowest, yielding Vowel Deletion, in the third, NOHIATUS is outranked by the other two constraints, the result being the preservation of the hiatus.

As for the off-glides, [u̯] appears again to be more vocalic than [i̯], in that it has a stronger tendency to be realized as a vowel, often being in free variation with [u]. Furthermore, when an /I/-/U/ or /U/-/I/ sequence occurs, /U/ is more often realized as a vowel and /I/ as a glide than vice-versa (e.g. *fort[ui̯]to* 'fortuitous', *p[ju]ttosto* 'rather'). However, another factor may counteract this tendency: since, according to a directional approach, Italian exhibits right-to-left syllabification (see 1.3), in a sequence of two high vocoids the rightmost tends to be stronger, often yielding a rising diphthong even when the sequence is /U/-/I/ (e.g. *eq[wi]no* 'equine').[22]

2.2.2 Glide alternations

The above-mentioned considerations do not affect the question of the phonological status of Italian glides. In fact, both on-glides exhibit a strong tendency to alternate with the corresponding N2-vocoids, as well as with the homorganic vowels. Moreover, diphthongisation is a widespread process in Italian.

2.2.2.1 Diphthongization

Italian displays a great variety of examples of both synchronic and diachronic diphthongization (cf. 2.1). During the 7th century, in fact, a process of rising diphthongization affected Latin stressed [e] and [o] which, through an intermediate stage of lengthening which yielded [ɛ:] and [ɔ:], became [jɛ] and [wɔ] respectively, as in *omo* > *uomo* 'man' or *petra* > *pietra* 'stone' (cf. Burzio 1994:28). The change, though, did not spread in a uniform fashion, so that some words were unaffected: this can often (but not always) be attributed to the position of the relevant vowel or to a preceding consonant cluster (Di Luzio 1988:302-9).

Among the affected roots, those which exhibit an alternation in the placement of stress have maintained the monophthong ~ diphthong alternation at the synchronic level, yielding a diphthongized form where the stress falls on the relevant syllable, a

[22] Harris and Kaisse (1999:135-6) point out that in Spanish, where the two high vocoids seem to have the same degree of sonority, this is the rule.

monophthongized form where the relevant syllable is unstressed. Examples are *p[j]éde* 'foot'/ *pedéstre* 'pedestrian', *s[w]óno* 'sound'/ *sonóro* 'sonorous'. Such alternating diphthongs are traditionally called *dittonghi mobili* 'movable diphthongs', and their distribution can also be conditioned by factors other than stress, as in some irregular verbs like *vóglio* 'I want'/ *v[w]ói* 'you want', *c[w]ócio* 'I cook'/ *cóssi* 'I cooked', *véngo* 'I come'/ *v[j]éni* 'you come' (cf. Di Luzio 1988:309).

On the other hand, there are cases of diachronic diphthongization which have become fossilized, thus yielding diphthongized allomorphs even when the relevant syllable is unstressed. *P[j]étra*, for instance, yields the unstressed forms *p[j]etróso* 'stony', and *p[j]etrificáre* 'to petrify', with the optional monophthongized forms *petróso* and *petrificáre* having become obsolete or poetic. In fact, Vincent (1987:276) remarks that the pattern of movable diphthongs is 'gradually eroded away by analogical generalizations in both directions (e.g. *suono* ~ *sonato* > *suonato*, *provare* ~ *pruovo* > *provo*)'.

As was mentioned in 2.2.1, contrary to on-glides that are not derived by diphthongization, glides in *dittonghi mobili* seem to be placed in the nucleus. This makes sense because, as Harris and Kaisse (1999:128) say as regards a similar phenomenon in Spanish, the mid vowels are nuclei when they break up, and there is no reason to stipulate an extra process by which the new-formed glides move out of the nucleus at some point in the derivation.

While diachronically it is diphthongization that took place, we can assume that synchronically we are facing a process of monophthongization in unstressed syllables. From an OT perspective, the constraint that militates against monophthongization is the faithfulness constraint UNIFORMITY-IO. On the other hand, since the glide + vowel sequence is part of a complex nucleus, the markedness constraint *COMPLEX$^{\text{NUC}}$ bans diphthongs. These, however, do appear in stressed syllables, and therefore *COMPLEX$^{\text{NUC}}$ must be split into *COMPLEX$^{\text{NUC}}$-STRESS and *COMPLEX$^{\text{NUC}}$-UNSTRESS: the first is dominated by UNIFORMITY-IO, while the second dominates it. Simplified tableaux for the words *piéde* and *pedéstre* are reported in (9a) and (9b) respectively.

(9) (a)

Input: /plɛde/	*COMPLEX^NUC-UNSTRESS	UNIFORMITY-IO	*COMPLEX^NUC-STRESS
a. ☞'pjɛde			*
b. 'pede		*!	

(b)

Input: /plɛdɛstre/	*COMPLEX^NUC-UNSTRESS	UNIFORMITY-IO	*COMPLEX^NUC-STRESS
a. pjɛ'dɛstre	*!		
b. ☞pe'dɛstre		*	

2.2.2.2 Glide Formation

As was mentioned above, alternation between glides and the homorganic high vowels is also a widespread phenomenon. Contrary to French and English, in Italian it applies not only inside (e.g. *t[u]* 'you'/ *t[w]oi* 'your (m.pl.)', *Mazzin[i]*/ *mazzin[j]ano* 'Mazzinian'), but also outside the PW (e.g. *I ragazz[i] sono usciti* 'The boys have gone out'/ *I ragazz[j] hanno*[23] *mangiato* 'The boys have eaten'). It can therefore be labelled as a connected speech process.

In Bosisio (1998) I have shown that, in a prosodic phonology perspective, so-called 'Glide Formation' in Italian appears to have the phonological utterance (PU) as its domain of application. As can be seen from the examples in (10), it applies not only between the two elements of a compound (10a), between a clitic and its host (10b), between two separate words forming a phonological phrase (10c), and between two words belonging to different phonological phrases within an IPh (10d), but also at the juncture between two distinct IPh within a PU (10e).

(10) (a) $[\text{Pulisc[j]}]_{PW}$ $[\text{orecchi}]_{PW}$

'Ear-pick'

(b) $[\text{M[j]}]_{PW}$ $[\text{hai}]_{PW}$ veramente delusa

'You have really disappointed me'

[23] In Italian, the grapheme <h> is not pronounced, therefore a word like *hanno* [anno] is vowel-initial.

(c) Sono [quas[j]]$_{CG}$ [arrivate]$_{CG}$

'They are almost there'

(d) [Niedd[w]]$_{PPh}$ [è partito]$_{PPh}$ per la Sardegna.

'Nieddu has left for Sardinia'

(e) [Giovann[j]]$_{IPh}$, [anche se ti vuole bene]$_{IPh}$, non lo permetterà.

'Giovanni, even if he loves you, won't allow this to happen'

Across PU boundaries, on the other hand, GF is blocked, as in (11):

(11) Partiamo doman[i]. A proposito, hai chiamato il taxi?

'We're leaving tomorrow. By the way, have you called a taxi?'

However, in a case like (12), GF appears to take place even if the two affected words belong to different PUs.

(12) Basta con i biscott[j]. Ho voglia di qualcosa di salato.

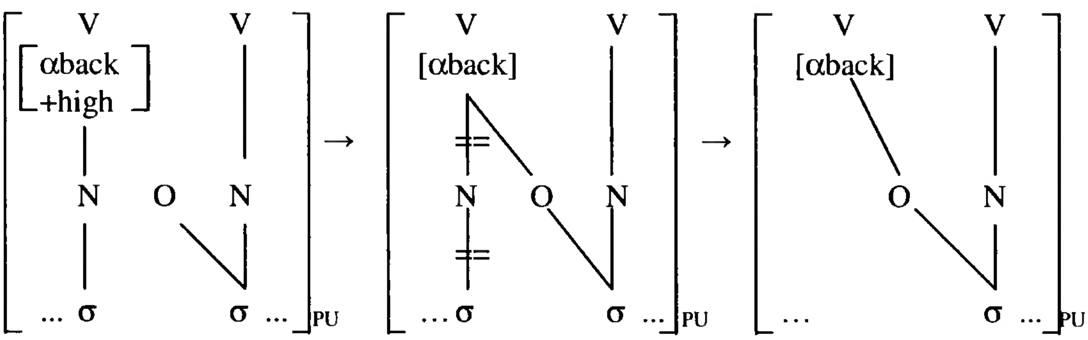
'Enough of biscuits. I feel like something salty'

It can be noted, though, that the two PUs are short and connected by a strong semantic tie. Moreover, GF can actually be blocked when they are pronounced at a slow tempo and a pause is inserted between them. This means that, at a normal speech rate – and when no pause intervenes – the two PUs are restructured into one, as in (13):

(13) [Basta con i biscott[j]. Ho voglia di qualcosa di salato]$_{PU}$

It can therefore be claimed that GF in Italian is a PU-span rule, and that it can provisionally be formulated as in (14).

(14) Glide Formation in Italian (provisional):



According to an OT interpretation, as was mentioned in 2.1, it is the constraint NoHiatus (or Onset) that triggers GF. In Italian, as in many other languages, such a constraint dominates those constraints that prevent high vowels from appearing in marginal syllabic positions (i.e. *M/U and *M/I), but is generally dominated by constraints militating against the appearance of non-high vowels in such positions (i.e. *M/a, *M/o, *M/e). This partial ranking is subsumed in (15), while the tableau in (16) gives evidence for it, based on the diverging behaviour of *tutte avevano* [tut.te.a.ve.va.no] 'everybody (f.) had' and *tutti avevano* [tut.tja.ve.va.no] 'everybody (m.) had'.

(15) *M/a, *M/o, *M/e » NoHiatus (Onset) » *M/I, *M/U.

(16) (a)

Input: /tUtte avevano/	*M/a, *M/o, *M/e	NoHiatus	*M/I, *M/U
a. tut.te̯a.ve.va.no	*!		
b. ☞ tut.te.a.ve.va.no		*	

(b)

Input: /tUttI avevano/	*M/a, *M/o, *M/e	NoHiatus	*M/I, *M/U
a. ☞ tut.tja.ve.va.no			*
b. tut.ti.a.ve.va.no		*!	

The domain of application of the process is determined by the alignment constraint Align-R (PU, σ), which requires that the right edge of a PU coincide with the right edge of a syllable, clearly militating against resyllabification across PU boundaries, and thus preventing GF in such a position. The ranking here will therefore be ALIGN-R (PU, σ) » NOHIATUS (ONSET). The latter, though, must be ranked above all other prosodic alignment constraints (ALIGN-R (IPh, σ), ALIGN-R (PPh, σ), etc.) in order to allow GF across IPh, PPh, CG, PW and F boundaries (see simplified tableaux in (17)). Notice that all such constraints must be explicitly ranked, since this notation does not involve the inclusion of any such constraint in any other. For instance, the requirement ALIGN-R (PPh, σ) does not imply ALIGN-R (IPh, σ), as instead strict layering and logical thinking would demand.

(17) (a)

Input: /domanI a/	ALIGN-R (UP, σ)	NOHIATUS	ALIGN-R (IPh, σ)
a. ☞ do.ma.ni.]$_{PU}$a		*	
b. do.ma.nj]$_{PU}$a	*!		*

(b)

Input: /dʒovannI aŋke/	ALIGN-R (UP, σ)	NOHIATUS	ALIGN-R (IPh, σ)
a. dʒo.van.ni.]$_{IPh}$aŋke		*!	
b. ☞ dʒo.van.nj]$_{IPh}$aŋke			*

2.2.2.2.1 Conditioning factors

It is worth spending a few more words on the conditioning factors to which glide ~ vowel alternations are subject. It must be borne in mind that GF, as a typical connected speech phenomenon, is an optional process, even if in fluent speech it is highly favoured. Its optionality is due mainly to the fact that it can be blocked by factors such as a slow speech rate (cf. Marotta 1987:876), the insertion of a pause or constituent-final lengthening, which in their turn are often conditioned not only by phonological, but also by semantic, pragmatic and even extra-linguistic information (cf. 1.2). Such factors as speech rate and pause insertion, however, can be considered as only indirectly relevant to connected speech phenomena: a different tempo can trigger restructuring of the domain of application of a phonological operation, especially when this is IPh or PU, so that the operation itself applies more or less often. In order to decide what the domain of application of a process is, it will thus be necessary to refer to the rather imprecise but naturally intuitive notion of 'normal speech rate' (Nespor and Vogel 1986:23-4).

A much more clearly and consistently determinant factor in blocking the formation of a glide in Italian is the position of stress. As in most other languages (see 2.1), GF never takes place when the clashing high vowel bears prominence. In other words, a stressed vocoid always occupies an N1 position, as can be seen in (18a) and (18b).

(18) (a) Non è poi cos[i] antipatico.
'He is not so unpleasant, after all'

(b) Vai t[u] a prenderla.
'*You* (sg.) go to pick her up'

In the two examples, the first clashing word – *così* 'so' and *tu* 'you' respectively – ends in a stressed high vocoid. This, being longer and louder, cannot lose its syllabicity, thus leaving the following onset empty.

This, however, is not the whole story, as a glance at the sentences in (19) and (20) will immediately reveal.

(19) (a) L'ho trovata determinata, quas[j] ostináta.
'I found her determined, almost stubborn'

(b) L'ho trovata difficile, quas[i] óstica.
'I found her difficult, almost fastidious'

(20) (a) Er[j] ancóra andata a prenderlo.
'You had even gone (fem.) to pick him up'

(b) Er[i] ánche andata a prenderlo.
'You had even gone (fem.) to pick him up'

In both (19a) and (19b), the first word involved in the clash is *quási* 'almost', ending in an unstressed high vowel. Nevertheless, while in the first instance GF regularly takes place, in the second one the process is blocked. Similarly, the unstressed high vocoid of *éri* surfaces as an on-glide in (20a), as a vowel in (20b). This phenomenon can be explained on the basis of PW[24]-stress placement on the second clashing word: whereas the initial vowel of *ostináta* 'stubborn' and *ancóra* 'even' (in (19a) and (20a) respectively) does not bear prominence, *óstica* 'fastidious' and *ánche* 'even' ((19b) and (20b)) have a stress on the clashing vocoid. This appears to indicate that, in Italian, GF is blocked not only when the first (the high) clashing vocoid is stressed, but also when the second one is. In other words, both vocoids have to be unstressed for GF to take place.

When the alternation takes place inside the PW, however, this condition does not seem to apply. Even if the high vocoid is still required to be unstressed, the other clashing vowel can – and often does – bear prominence (e.g. *berluscon[j]áno* 'Berlusconian'). This appears to indicate that two different processes of glide ~ vowel alternation take place in Italian: one at the PW level and one at the PU level, i.e. one inside and one outside the PW. While GF at the PU level is a postlexical process, the formation of a glide within PWs applies in the lexicon, as a direct consequence of initial syllabification (see below). The reason why the second process is unaffected by stress placement is thus a natural consequence of the fact that, in languages like Italian and Spanish (cf. Harris and Kaisse 1999:124), stress assignment must follow initial syllabification, since it depends on syllable count and, partially, on syllable weight. As for the absence of stress on the high vocoid, it is a consequence, rather than a cause, of its being syllabified as a glide (see 2.2.3 below).

In OT, as was said in 2.1, high ranking of the constraint PK-MAX-IO – by which if a segment is the stress peak of the input its correspondent must be the stress peak of the output – accounts for GF blocking when the (first) high vowel is stressed. In a language like Italian, therefore, PK-MAX-IO must dominate NOHIATUS, as is shown in the (very simplified) tableau of *cosí amici* [ko'zia'mi:tʃi] 'such friends, so friendly', where the square brackets indicate foot structure.

(21)

Input: /kosI amItʃI/	PK-MAX-IO	PARSE-SYL	NOHIATUS	*M/I
a. ☞ ko[zia][mi:tʃi]		*	*	
b. [kozja][mi:tʃi]	*!			*

[24] Notice that, while secondary stress on the first vocoid prevents GF form applying (e.g. *tiamina* [ˌti.a.'mi:.na] 'thiamine'), stress on the second vocoid, in order to block GF, must be main word stress.

The tableau in (21) shows that PK-MAX-IO does not only dominate NOHIATUS, but also PARSE-SYL. The ranking of the latter with respect to NOHIATUS is irrelevant here. However, these observations alone cannot account for GF blocking when the second vowel, rather than the first, is stressed. The ranking of other constraints must be taken into consideration, such as FT-BIN and the more specific ANCHORING-IO(PW, L)-STRESS, which requires every segment at the left periphery of the input PW[25] that bears PW stress to have a correspondent at the left periphery of the output PW, thus preventing resyllabification before a stressed vowel.

The tableau in (22) shows that the ranking PK-MAX-IO » ANCHORING-IO(PW, L)-STRESS (abbreviated to ANCHOR-IO-STRESS) » NOHIATUS, PARSE-SYL » *M/I yields the correct output of *téneri ánimi* ['tɛneri'animi] 'tender souls, inclinations', where GF is blocked by PW stress on the second clashing vocoid.[26]

(22)

Input: / tɛnerI anImI/	PK-MAX-IO	ANCHOR-IO-STRESS	NOHIATUS	PARSE-SYL	*M/I
a. ☞ [tɛneri][animi]			*		
b.[tɛnerja][nimi]	*!	*			*
c.[tɛneri]a[nimi]	*!	*	*	*	
d.[tɛne][rjani]mi		*!		*	*
e.[tɛne][ria][nimi]	*!	*	*		
f.[tɛne]ri[ani]mi			*	*!*	

(23) and (24) report the tableaux for the phrases *téneri amíci* ['tɛnerja'mitʃi] 'tender friends' and *téneri animáli* ['tɛne,rjani'ma:li] 'tender animals', in which the second clashing vocoid is either unstressed or bears only secondary stress. The tableaux show how, when the second vocoid does not bear PW prominence, ANCHORING-IO(PW, L)-STRESS remains vacuously unviolated, thus leaving NOHIATUS and PARSE-SYL to select the most harmonic output. Since the winning candidates do not violate either, these two constraints cannot be ranked with respect to each other.[27]

---

[25] Reference to the PW is important here, in that, within the PW, on-glides easily occur before stressed vocoids (cf. *berluscon[j]áno* 'Berlusconian').

[26] Notice, however, that PK-MAX-IO still plays an important part in selecting the winning candidate.

[27] Since the evidence for its ranking is not conclusive and irrelevant to our purposes, the constraint FT-BIN has been left out. However, its location in the constraint hierarchy is not without consequences: in tableau (23), for instance, if FT-BIN were to be ranked above PARSE-SYL the winning candidate would be b. rather than f. Notice that, in such a case, the ranking of NOHIATUS with respect to PARSE-SYL would also become relevant.

(23)

Input: /tɛnerI amItʃI/	PK-MAX-IO	ANCHOR-IO-STRESS	NOHIATUS	PARSE-SYL	*M/I
a.[tɛneri]a[mitʃi]			*!	*	
b.[tɛne]rja[mitʃi]				*!	*
c.[tɛneri][amitʃi]	*!		*		
d.[tɛne][ria][mitʃi]			*!		
e.[tɛne][rjami]tʃi	*!			*	*
f. ☞[tɛnerja][mitʃi]					*

(24)

Input: /tɛnerI anImalI/	PK-MAX-IO	ANCHOR-IO-STRESS	NOHIATUS	PARSE-SYL	*M/I
a. ☞[tɛne][rjani][mali]					*
b.[tɛneri][ani][mali]			*!		
c.[tɛne]ri[ani][mali]			*!	*	
d.[tɛne][ria][nima]li	*!		*	*	
e.[tɛneri]a[nimali]	*!		*	*	
f.[tɛnerja][nimali]	*!				*

The examples illustrate how, in an OT framework, the use of anchoring constraints making explicit reference to the PW obviates the need to specify that on-glides may precede stressed vowels within such consituents.

2.2.2.2.2 Syllabification and resyllabification

As was mentioned in section 2.1, glide ~ vowel alternations are closely connected to syllable structuring: according to the environment in which it is placed, a vocoid can occupy an O, an N1, an N2, an N3 or a Co[28] position within the syllable.

It was also pointed out (cf. 1.3) that languages seem to behave in different ways with respect to syllabification, i.e. the parsing of segments into syllables. Italian probably undergoes an initial stage of syllabification, which takes place at the lexical level. As the other Romance languages, moreover, it appears to exhibit a second, postlexical

[28] As was said above, however, Italian does not seem to have Co-glides: vocoids in this language can only be linked to four syllabic positions.

process of resyllabification, by which syllables are restructured even across word boundaries.[29] An Italian phrase like (25a), for example, will be typically resyllabified as in (25b).

(25) (a) Con un amico
'With a friend (masc.)'
(b) [kɔ.nu.na.mi.ko]

By looking at the way Syntactic Doubling (a process of Standard Italian by which a word-initial consonant is geminated when preceded by a word carrying the main stress on its final vowel) patterns, Nespor and Vogel (1986:67-72) argue that the domain for resyllabification in Italian is PPh, as it is in French. The fact that GF applies across IPh boundaries, however, requires resyllabification to apply in a larger domain, probably the PU. Just as GF within the PW is the effect of lexical syllabification, GF in larger domains evidently results from resyllabification at the phrase or utterance level. It is not by chance that Germanic languages, in which resyllabification does not generally take place, tend to lack GF outside the PW.

In an OT framework, the constraint whose ranking regulates the application of syllabification – and thus of GF - across PW boundaries is ALIGN-R(PW, σ), which calls for the coincidence between the right edge of a PW and that of a syllable. In Germanic-type languages, this constraint dominates NOHIATUS, while in Romance-type languages it is dominated by it. The contrast between the two different rankings is shown in tableaux (26) and (27). These illustrate, respectively, the evaluation of the Italian phrase *molti attori* ['mol.tjat.'to.ri] 'many actors', in which GF (and resyllabification) takes place, and of the English phrase *very annoyed* ['ve.ri.jə.'nɔi̯d], in which GF is replaced by Glide Insertion (GI), which does not involve resyllabification. As will be shown in more detail in chapter 5, GI is triggered by the ranking NOHIATUS » DEP-C-IO.

[29] Besides, as was mentioned in 1.3, in Italian there are segments that seem to be already assigned a syllabic location underlyingly (also see 2.2.3 below). This observation can be captured either by means of partial underlying syllabification - i.e. by pre-linking those segments to specific positions (e.g. rhyme heads) – or, less problematically, by assigning them C or V positions on a CVX-skeleton.

(26)

Input: /moltI attorI/	NOHIATUS	DEP-C-IO[30]	ALIGN-R(PW, σ)
a. ☞ mol.tjat.to.ri			*
b. mol.ti.at.to.ri	*!		
c. mol.ti.jat.to.ri		*!	

(27)

Input: /verI ənɔId /	ALIGN-R(PW, σ)	NOHIATUS	DEP-C-IO
a. ve.rjə.nɔ̯id	*!		
b. ve.ri.ə.nɔ̯Id		*!	
c. ☞ ve.ri.jə.nɔ̯id			*

2.2.3 The phonological status of Italian glides

From what has been said, it is clear that the overwhelming majority of glides in Italian appear to be allophones of the homorganic vowels rather than forming independent phonemes. The question is, however, are there any cases in which a distinction must be made between a glide and the corresponding high vowel underlyingly? This is not an easy question to answer. It is sometimes even difficult to decide whether a given word contains a glide or a vowel, because things may change according to the speaker or to the situation. Marotta (1987:879-81), for example, states that *arguire* 'to deduce', *arcuato* 'bent' and *acuire* 'to sharpen' are pronounced with an [u], while in my Italian such words definitely contain a [w]. Moreover, the absence in Italian of the sequences [ji] and [wu] deprives us of a possible clue to the dilemma of the phonological opposition between glides and high vowels (ibid.:857).

Serianni (1989:20) actually argues that there exist some minimal pairs, such as *allev[j]amo* 'we breed' and *allev[i]amo* 'we relieve', *sp[j]anti* 'you (sg.) dig out' and *sp[i]anti* 'spying', *la q[w]ale* 'which (fem.sg.)' and *lac[u]ale* 'lacustrine'. In all these cases, however, the second item can also be pronounced with a glide. Moreover, it can be argued that a word like *sp[i]anti* does not exhibit a glide because its high vocoid carries over a secondary stress in its derivation from *spía* 'spy (n.)'. The pair *allev[j]amo/ allev[i]amo* can also be ruled out on the basis of the fact that,

---

[30] Since GI does apply in certain cases in Italian, DEP-C-IO must be ranked below NOHIATUS in this language, too.

contrary to *allev[j]amo* (*allev+iamo*), the second word (*allevi+iamo*) contains a high front vocoid not only in the suffix, but also in the root. The presence of a vowel in *lacuale* (*lacu+ale*), on the other hand, cannot be explained by invoking the notion of morpheme boundary, because there are examples of glides regularly surfacing before the same suffix *–ale* (cf. *port[w]ale* 'dock worker').

A plausible diagnostic for finding underlying glides is their presence in GF blocking contexts. According to Marotta (1987:875) in Italian, as in French, GF tends to be blocked by a preceding obstruent + liquid (OL) cluster, since the sequence OL-glide is forbidden inside the syllable for reasons of sonority distance. Before the same suffix *-ábile*, for example, we find *fr[i]abile* 'crumbly' as opposed to *encom[j]abile* 'praiseworthy', and words like *triángolo* 'triangle', *cliénte* 'customer', *Adriáno* 'Hadrian', *cóngruo* 'congruous' all contain a full high vowel. Other lexical items, such as *patr[j]a* 'home-country', *fl[w]oro* 'fluorine', *tr[w]ogolo* 'trough' or *industr[j]a* 'industry', on the other hand, normally have a glide, even if the stylistic or pragmatic context may trigger the production of a hiatus (cf. Marotta 1987:875). A possible explanation is that, in these latter cases, the glide is already present underlyingly. In CVX-tier terms, such vocoids should then be pre-linked to a C slot. Marotta (ibid.), however, points out that, if the high vowel belongs to a suffix (together with the following vowel), GF generally takes place even after an OL cluster (cf. *destr[j]ero* 'steed', *scopr[j]amo* 'we find out', *compr[j]ate* 'you (pl.) buy (subj.)'), although GF blocking regularly takes place when the morpheme boundary follows the high vocoid (cf. *pr[i]+ore* 'prior' or *obl[i]+are*[31] 'to forget'). The word *industria* may well be included in the first category, since it appears to be formed by a root *industr-* and a suffix *–ia* (cf. *industr-e* 'industrious', from the same root). Similar considerations can be made as regards the word *patria*, in which the suffix –*ia* has been combined with the Latin root *patr-* (from *pater/patris* 'father'), no longer productive in Italian. An example such as *truogolo* is also suspicious, because its glide was diachronically derived by means of diphthongization and thus may be in the nucleus rather than in the onset. The same, however, cannot be said of *fluoro*, deriving from Latin *fluere*.

---

[31] Furthermore, *obl[i]are,* just like *sp[i]are* 'to spy' (see above), derives from a form (*oblío* 'forgetfulness') in which the high vowel is stressed.

In a nutshell, the evidence in favour of a phonological distinction between glides and vowels in Italian is by no means conclusive, also due to the fact that glides and homorganic vowels are often in free variation. Words such as *fluoro* 'fluorine', however, must be marked as exceptions by pre-linking the high vocoid to a C position on the skeleton (28).

(28)

f	l	U	ɔ	r	o
C	C	C	V	C	V

By the same token, if one considers *lacuale* 'lacustrine' as constantly exhibiting a vowel at a normal speech rate, while surfacing with a glide only at a fast rate, its /U/ must be pre-linked to a V position underlyingly (29).

(29)

l	a	k	U	a	l	e
C	V	C	V	V	C	V

As was shown in 1.3, the [i] in Italian words like *fantasia* [ˌfan.ta.'zi.a] 'imagination' can also be represented as pre-linked to a V position on the skeleton. An alternative interpretation may be that, since the vowel bears main word prominence, such a stress could be already present underlyingly, thus preventing PW-level GF. However, as was mentioned above, in Italian initial syllabification precedes stress assignment, and thus stress placement cannot affect the syllabicity of segments word-internally. In Harris and Kaisse's (1999:24) words, the fact that stressed high vocoids do not form glides within words is due to 'a secondary effect that results from contrastive syllabicity': syllabicity, and not stress, is lexically contrastive here.

As was said in 2.1, the problem of the phonological status of glides is hardly relevant to OT, because of the Richness of the Base. The above-mentioned examples, however, can be subjected to an OT analysis. The fact that GF is blocked after an

OL-cluster is easily accounted for by means of the constraint *OLG, which must be ranked above NoHiatus. When a morpheme boundary intervenes between the OL-cluster and the high vocoid, however, GF does apply. The *OLG constraint must thus be modified to include the specification of such a boundary, yielding *OLG]$_m$ (no OLG sequences within the same morpheme). The tableaux for *cliente* [kli.ɛn.te] 'customer', where the OL-cluster is contained within the same morpheme as the high vocoid, and *destriero* [des.trjɛ.ro] 'steed', in which a morpheme boundary separates the high vocoid from the OL-cluster, are reported in (30) and (31), respectively.

(30)

Input: /klIɛnte/	*OLG]$_m$	NoHiatus
a. ☞ kli.ɛn.te		*
b. kljɛn.te	*!	

(31)

Input: /destr + Iɛro/	*OLG]$_m$	NoHiatus
a. des.tri.ɛ.ro		*!
b. ☞ des.trjɛ.ro		

Explaining why words such as *fluoro* [flwɔro] 'fluorine' and *truogolo* [trwɔgɔlo] 'trough' contain an OLG cluster, however, is problematic within an OT framework. In such forms, the consonantality of the glide should be somehow marked in the input as a lexical exception, thus undermining the Richness of the Base priniciple. The same can be said of the high vowel of *lacuale* 'lacustrine'.

2.3 Alternations in Ammani Arabic

In this section I will tackle the issue of glide ~ vowel alternations in Ammani Arabic (AA). I will first discuss what segments can be construed as glides in this variety, then I will show how vocoids can be represented in a non-concatenative language and what positions they may occupy in the AA syllable. Next, I will deal with the syllabic alternations of the vocoids, their domain of application and the factors conditioning them. Finally, I will try to shed some light on the difficult question of the phonological status of glides in AA and in Arabic at large. For each issue an alternative OT interpretation will be given.

2.3.1 Arabic glides and the problem of gutturals

Apart from the two prototypical glides [j] and [w], Arabic exhibits other sounds which have at times been classified as glides, i.e. the glottals [ʔ] and [h] and the pharyngeals [ʕ] and [ħ] (cf. McCarthy 1981).

Chomsky and Halle (1968), based on Jakobson, Fant & Halle's (1952) observations of the peculiar spectra of these sounds, include [ʔ] and [h] in the category of glides, and assign them the features [+sonorant, -consonantal, -vocalic][32]. As evidence that [h] should be classified as a glide, Andersen (1972) analyzes the process of 'peak attenuation' (V > h + V) as a type of diphthongization, considering the inverse process (h + V > V) as a monophthongization. In certain languages, among which Malay (see 5.1) and the Kalinga dialect of the Filipino language Guininaag (Gieser 1970, in Kenstowicz and Kisseberth 1979:244), [ʔ] alternates with [j] and [w] as a hiatus breaker. In the latter variety, for instance, [j] is inserted when the first of two clashing vowels is /i/ (32a), [w] when it is /u/ (32b) and [ʔ] when it is /a/ (32c).

(32) (a) *daŋli* 'believing' → *daŋlijan* 'to believe'

(b) *ʔudugu* 'pity' → *ʔuduguwan* 'to pity'

(c) *ʔala* 'getting' → *ʔalaʔan* 'to get'

(Kenstowicz and Kisseberth 1979:244)

Kreidler (1989:68) – following Ladefoged (1962) - points out that [h] can be seen as the voiceless equivalent of whatever vowel comes next, and should therefore be

[32] Schane (1973:27), on the other hand, draws a distinction between what he terms 'laryngeal glides' ([ʔ] and [h]), classified as [-syllabic, -consonantal, -sonorant] and semivowels, classified as [-syllabic, -consonantal, +sonorant].

classified as a semivowel. Ladefoged and Maddieson (1996:325-6), however, argue that this simply means that [h] - and its voiced counterpart [ɦ] – only exhibit a laryngeal specification, assuming whatever supralaryngeal features are provided by the surrounding sounds. Moreover, these authors (ibid.:326) remark that some languages - among which are, interestingly, Hebrew and Arabic (Laufer 1991) – 'show a more definite displacement of the formant frequencies for **h**, suggesting it has a constriction associated with its production.'

Other authors (e.g. Pike 1943, Gimson 1965, Abercrombie 1967, Lass 1976, Durand 1987, Angoujard 1990) are even clearer in denying glottals any status as 'glides' (and hence as 'vocoids'), at least from a phonological point of view. Lass (1976: 145-167) builds his argument especially on the notion that non-oral – i.e. laryngeal – strictures are true consonantal articulations: sounds like [h] and [ʔ] being pulmonic egressives, 'a glottal closure [...] is simply the first in a series of possible obstructions of the air-stream' (ibid.:148). Hence, [ʔ] can be easily classified as a (de-oralized) stop, [h] as a (de-oralized) fricative. Lass (ibid.:148-50) then puts forward a more strictly phonological argument: in those varieties of English in which [ʔ] replaces some other segments, these are generally [-son, -cont, -voice]. The choice of a 'glide' – and of this particular 'glide' – to replace a voiceless stop would be difficult to explain, considering the features involved. Lass (ibid.:159-163) – although admitting that the synchronic evidence is rather scanty[33] – also produces substantial historical evidence for the interpretation of [h] as a non-oral fricative.

Durand (1987), drawing evidence from different phonological processes, shows that in Malay [ʔ] and [h] are best treated as obstruents and suggests that such considerations should be extended to other languages. For instance, in Malay a stem-initial voiceless obstuent undergoes deletion when preceded by a prefix-final nasal, after this latter has assimilated in place of articulation with such an obstruent (33a). As can be seen in (33b), [h] and [ʔ] behave exactly like voiceless obstruents, while the sonorants /j/, /w/, /l/ and /r/ are not deleted, but actually trigger the deletion of a preceding prefix-final nasal (33c).

[33] However, a case like that of the Florentine vernacular, in which t ~ θ, p ~ ϕ and k ~ h, offers a clear parallel between [h] and the other voiceless fricatives.

(33) (a)	/məŋ=tulis/	[mənules]	'to write (act.)'
	/məŋ=pukul/	[məmukol]	'to beat (act.)'
	/məŋ =kawal/	[məŋawal]	'to guard (act.)'
(b)	/məŋ =hakis/	[məŋakes]	'to erode (act.)'
	/məŋ=hantar/	[məŋantar]	'to send (act.)'[34]
(c)	/məŋ =ratʃun/	[məratʃon]	'to poison (act.)'
	/ məŋ =lawat/	[məlawat]	'to visit (act.)'
	/ məŋ=wakili/	[məwakili]	'to represent (act.)'
	/ məŋ=jakini/	[məjakini]	'to be convinced (act.)'
			(from Durand 1987:83-6)

Another process that supports the notion of Malay [h] and [ʔ] as obstruents rather than glides is the syllable-final weakening of /k/ and /s/ to [ʔ] and [h] respectively (34).

(34)	/masak/	[masa**ʔ**]	'to cook'
	/masakan/	[masakan]	'the cooking'
	/masak+kan/	[masa**ʔ**kan]	'to cause to cook for'
	/kipas/	[kipas ~ kipa**h**] [35]	'fan'
	/kipas+kan/	[kipaskan ~ kipa**h**kan]	'to cause to fan for'
	/pən=kipas+an/	[pəŋipasan]	'the fanning of'
			(ibid.:87)

As far as Arabic is concerned, several scholars (e.g. Abdo 1969, Abumdas 1985, Watson 1989) classify [ʔ] and [h] as glides. Abumdas (1985:168-171), for instance, points out that in Libyan Arabic [ʔ] and [h], just like [j] and [w], undergo deletion in

[34] Durand (1987:84) does not give any examples with /ʔ/ because he assumes that there is no such phoneme in Malay: contrary to [h], this sound appears only in epenthetic contexts.
[35] While the weakening of /k/ is obligatory, that of /s/ is optional.

intervocalic position (e.g. *naaD* < *nahaD* 'he arose', *raaf* < *raʔaf* 'he showed mercy', just like *ramaa* < *ramaja* 'to throw')[36] and monophthongization (e.g. *raas* < *raʔs* 'head' just like *beet* < *bayt* 'house', see below). This is not, however, the case in all Arabic dialects; in AA, for instance, intervocalic deletion[37] applies only to [j] and [w] (see 4.3), which shows that this kind of evidence is not conclusive.

Along the same lines, Watson (1989:163-165) claims that, in the Yemeni dialects she studied, glottals behave like (Co-)glides. In these varieties, in utterance-final position the long vowels /u:/ and /i:/ undergo a process of diphthongisation by which their second segment becomes a Co-glide - [w] and [j] respectively (35a-b). Since the non-high long vowel /a:/ becomes [ah] or [aʔ] in exactly the same environment (35c), Watson (ibid.) concludes that the two glottal sounds [h] and [ʔ] belong to the class of glides.

(35)	(a)	/simiʕ+uu/	→	[simiʕuw] $]_{PU}$ 'they (m.) heard'
	(b)	/simiʕ+nii/	→	[simiʕnij] $]_{PU}$ 'he heard me'
	(c)	/simiʕ+naa/	→	[simiʕnaʔ/h] $]_{PU}$ 'we heard'

(from Watson 1989:163-4)

In AA a similar phenomenon takes place, although its application appears to be more sporadic than in the Yemeni dialects. As is shown in (36), utterance-final /a:/ can be pronounced either as [a:] or [aʔ], utterance-final /u:/ surfaces as either [u:] or [uw], and utterance-final /i:/ is realized as either [i:] or [ij].

(36)	(a)	laa	→	l[a:] ~ l[aʔ] 'no'
	(b)	haada-l ᶜaduu	→	haada-l ᶜad[u:] ~ ᶜad[uw] 'this is the enemy'
	(c)	maa fii	→	maa f[i:] ~ f[ij] 'there isn't (any)'

In an OT perspective, PU-final diphthongization may be given by the combination of two markedness constraints: context-free *V: (see 2.3.3.1), requiring that no vowel be

[36] See 2.3.5 and 4.3 for more information about intervocalic Glide Deletion (GD).

long, and contextual *V]$_{PU}$ (see 3.3), by which a PU cannot end in a short vowel (i.e. in a monomoraic syllable). As is shown in tableau (37), both these constraints must dominate *FISSION, the constraint militating against diphthongization. Hence, the optimal candidate is diphthongized c.

(37)

Input: /fI:/	*V:	*V]$_{PU}$	*FISSION
a. fi:]$_{PU}$	*!		
b. fi]$_{PU}$		*!	
c. ☞ fij]$_{PU}$			*

If [ʔ] is seen as a glide, the same interpretation can be given for *l[a:] ~ l[aʔ]*; if instead it is considered a default epenthetic consonant, *V: and *V]$_{PU}$ must also dominate DEP-C-IO, the constraint militating against Consonant Insertion.

Other authors, in fact, do not believe there is sufficient ground to include glottals in the category of glides. Angoujard (1990:16), for instance, defines [ʔ] as an obstruent because of its low degree of sonority compared to the glides [w] and [j]. ʔ-insertion in Arabic, moreover, does not seem to be in complementary distribution with Glide Insertion (GI, see ch. 5). [38] In all varieties of Arabic, a [ʔ] is often[39] inserted in domain-initial position in order to avoid onsetless syllables (see 1.3).[40] The process

[37] Monophthongization, however, applies to AA glottals, too (e.g. /il-raʔi/ > [ir.raai] (Jordanian newspaper)).

[38] From a phonetic point of view, however, Al-Ani (1970:60) observes that initial [ʔ] is often realized 'as a short onset glide', especially when preceding /i/ or /i:/.

[39] This is an optional process, because so is resyllabification – with which ʔ-insertion is in complementary distribution - in AA and in similar dialects: ʔ-insertion (i) only applies when resyllabification has not provided an alternative (ii) to fill in the empty onset.

(i) /fi:/ + /l/ + /imtiħa:na:t/ → [fil.ʔim.ti.ħa.na:t] 'in the examinations' (no resyllabification)
(ii) /fi:/ + /l/ + /imtiħa:na:t/ → [fi.lim.ti.ħa:.na:t] 'in the examinations' (resyllabification)

[40] Some scholars (e.g. Harrell 1957) consider these word-initial [ʔ]s to be phonemic, and not inserted. They therefore postulate a process of ʔ-deletion to account for the cases in which such a phone does not surface. However, it may be observed that, at least in dialects like AA and Libyan Arabic (see Abumdas 1985:165), initial [ʔ] only appears in those derived forms which would otherwise begin with a vowel: when a derivational alternant from the same root does not begin with a vowel, the [ʔ] disappears (e.g. [ʔiktib] 'write!' vs [katab] 'he wrote'; [ʔabjaD] 'white (m.s.)' vs [beeDa] 'white (f.s.)'). To my knowledge, there are no AA words to which these considerations do not apply, which goes to show that initial [ʔ] is not a radical, but an epenthetic consonant. Moreover, the [ʔ] that, in varieties like AA, is derived from SA [q], and is therefore part of the root, is never deleted, which testifies to the different status of the two types of [ʔ]. Concerning words such as *(')uxt* 'sister' or *(')arD* 'earth', however, Cantineau (1960:118) remarks that speakers appear to perceive the presence of a *hamza* even when this

applies not only before /a/, but also before /i/ and /u/, and thus it is not paralleled by GI. Therefore, [ʔ] could be construed not so much as a glide, but rather as a default epenthetic consonant, which is inserted in domain-initial and in domain-final position in order to fill in an empty onset or an empty coda, respectively. Moreover, [ʔ] and [h] do not alternate with [a] as [j] alternates with [i] and [w] alternates with [u] (see below).

As for the pharyngeal – or, according to Ladefoged and Maddieson (1996), 'epiglottal' – sounds of Arabic, they have often been described as fricatives (see for instance Harrell 1957, Abumdas 1985, Durand 1987, Ladefoged and Maddieson 1996). In casual colloquial Arabic, however, they appear to be mostly pronounced as approximants (Ladefoged and Maddieson 1996:168, following Catford 1977). From a phonetic perspective, Ladefoged and Maddieson (ibid.) observe that, while the sound [ħ] is accompanied by audible local turbulence, it is rarely so for the sound [ʕ]. From a phonological point of view, however, the status of such sounds as glides is even more debatable than in the case of [h] and [ʔ].[41] In Arabic, as in other Semitic languages, these latter often pattern together with [ħ] and [ʕ], with which they form the traditional class of *gutturals*. These sounds take part in a number of phonological processes, which appear to set them apart from the glides [j] and [w] (Kenstowicz and Kisseberth 1979:244). In SA, for instance, /i/ is lowered to /a/ when in the environment of a pharyngeal or glottal consonant (ibid.:250).

On the other hand, [j] and [w] undergo (or are involved in) synchronic and diachronic processes – such as Glide Metathesis[42] (Brame 1970:293, Abumdas 1985:115), Bidirectional Assimilation of [w] to [j][43] (Abumdas 1985:140) and other [j] ~ [w] alternations[44] (Abumdas 1985:147-151), Vocalic Assimilation[45] (Brame 1970:97-111;

---

is not pronounced, thus considering such roots as trilateral rather than biliteral. In these cases, the [ʔ] may be present in the UR (or input).

[41] It has been noticed, though, that sometimes the pharyngeal [ʕ] seems to alternate with [a], just like [w] alternates with [u] and [j] with [i] (cf. Cantineau 1960:229 on the dialect of El-Hamma of Gabes).

[42] A process by which 'the sequence CGV changes to CVG then to CV: (Abumdas 1985:115), and which is common to several varieties, among which AA.

[43] A prevalently diachronic phenomenon assimilating [w] to [j] in either direction (Abumdas 1985:140).

[44] Abumdas (1985:173-179) gives examples of sporadic diachronic change from [ʔ] to glide as well, which would appear to show that [ʔ] actually does behave like a glide. In AA, however, only [w] and [j] seem to be involved in the alternation.

[45] A process by which the low vowel /a/ turns into a high ([u] or [i]) or mid ([o] or [e]) vowel when followed by a glide ([w] or [j]).

Abumdas 1985:154-5), monophthongization and glide ~ vowel alternations (see below) – which do not affect any other non-syllabic segments, including gutturals. Besides, weak adjectival roots, i.e. those with a comparative of the form aCCa (e.g. *ghaali* 'expensive' ~ *aghla* 'more expensive', *Hilw* 'sweet' ~ *aHla* 'sweeter') end in a vowel, [j] or [w]; never in a guttural.

Finally, in most – perhaps all - varieties of Arabic, the distribution of [w] and [j] differs from that of all other 'consonants', including gutturals: while all other non-syllabic sounds can appear in all positions with respect to adjacent vowels, there are some restrictions on the distribution of [j] and [w] (see below).

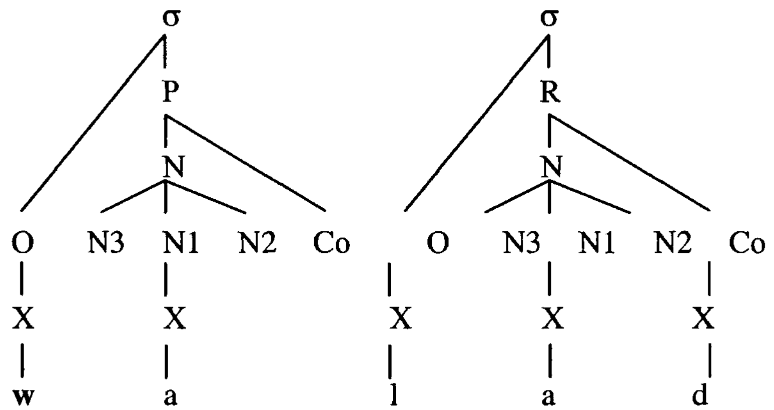
In conclusion, the question of the status of glottals and pharyngeals in Arabic – as in many other languages - appears to remain open. Their belonging to the class of glides is, to say the least, moot. Thus this section – as the rest of the thesis - will only deal with the two Arabic segments that can uncontroversially be classified as glides (and hence as vocoids), namely [j] and [w].

2.3.2 Distribution of Arabic high vocoids

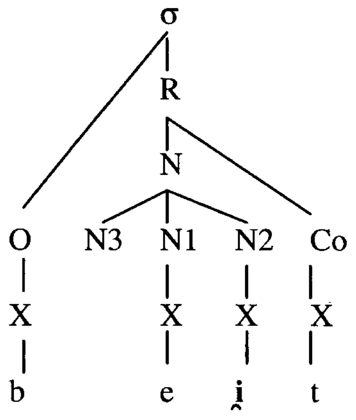
2.3.2.1 Distribution within the syllable

In SA, high vocoids can appear both as on-glides (38a) and as off-glides (38b).

(38) (a) *walad* 'boy'



(b) *bayt* 'house'



However in some dialects,[46] among which AA, a process of monophthongization often takes place to reduce /ai̯/ and /au̯/ clusters to the long vowels [e:] and [o:] respectively. A word like *zayt* 'oil' (SA [zei̯t]) is thus pronounced as [ze:t], a word like *mawz* 'banana' (SA [mau̯z]) as [mo:z]. Such a process appears to be obligatory only in closed syllables. Elsewhere, its application seems to vary from one lexical item to another: in words such as *awlaad* [ʔau̯la:d] 'children, boys' or *law* [lau̯] 'if' it does not apply at all; the pronunciation of other words - such as *dawla* 'state' – appears to vary according to the formality of the register, although it might be argued that, for the educated speaker, adopting a more formal register corresponds to code-switching to SA. In any case, it is clear that the URs of words like *bayt* [be:t] 'house' and *lawn* [lo:n] 'colour' are /baIt/ and /laUn/ respectively, since the high vocoid surfaces in the plurals *biyuut* [biju:t] and *alwaan* [ʔalwa:n].[47] Hence, even in varieties like AA, diphthongs like /aI/ and /aU/ can be found at least underlyingly, although a phonotactic constraint must be placed against the tautosyllabic sequence VGC to account for the obligatory monophthongization in closed syllables (also see Abumdas 1985:40).

Such a constraint is readily formulated in OT terms as $*VGC]_{\sigma}$. As shown in tableau (39), this is in conflict with the anti-monophthongization constraint UNIFORMITY-IO,

---

[46] For Egyptian Arabic, see Harrell (1957:38-41), for Libyan Arabic, see Abumdas (1985:40).

[47] Some authors, disregarding alternations such as *b[e:]t* 'house' ~ *bu[j]uut* 'houses', consider /e:/ and /o:/ to be independent phonemes (cf. Cantineau (1960:217) for the dialect of El-Hamma of Gabes (Tunisia), Harrell (1957:52) for Egyptian Arabic, Suleiman (1985:28-9) and Rammuny (1966) for Jordanian Arabic).

and must outrank it in order to obtain the correct output. Notice that *V:, too, must be outranked by *VGC]$_\sigma$, since monophthongization generates a long vowel. Moreover, in order to exclude the suboptimal candidate c., *V: must also be dominated by MAX-μ-IO (or MAX-X-IO)[48], which requires a mora (or a skeletal slot) in the input to have a correspondent in the output, thus militating against shortening. The ranking of MAX-μ-IO (MAX-X-IO) with respect to *VGC]$_\sigma$ and of UNIFORMITY-IO with respect to both MAX-μ-IO (MAX-X-IO) and *V: are irrelevant to this discussion.

(39)

Input: /laUn/	*VGC]$_\sigma$	MAX-μ-IO (MAX-X-IO)	UNIFORMITY-IO	*V:
a. lau̯n	*!			
b. ☞ lo:n			*	*
c. lon		*!	*	

As for the distribution of AA on-glides, they seem to follow a similar pattern to those of Spanish (see Harris and Kaisse 1999, in 2.1 and 2.2), i.e. they are probably syllabified in the onset when syllable-initial, in the nucleus when preceded by a tautosyllabic consonant. The fact itself that they may be preceded by a tautosyllabic consonant – as in, for instance, *kwejjes* 'good, well' – proves that, in a case like this, they are not onset segments, since in AA complex onsets are banned (see 1.3). On the other hand, the doubling of intervocalic glides (e.g. *awwal* 'first') is analogous to that of consonants, leading one to assume that the second element of such long segments may make up a syllable onset.

A distinction must also be drawn between off-glides (N2-glides), which are placed in the nucleus, and coda-glides (Co-glides), although the latter are not found underlyingly but are exclusively the result of the process of diphthongization that was mentioned in 2.3.1 (see Watson 1989).

In many dialects, restrictions on the distribution of glides with respect to their quality have been pointed out. For example, Al-Ani (1970:35) notes that, in Iraqi SA, the only possible (falling) diphthongs are [ai̯] and [au̯]. This seems to be true for most Arabic dialects, although Harrell (1957:38-41) remarks that, in Egyptian Arabic, the

48 In cases like this, in which only nuclear (and thus moraic) vocoids are involved, the use of either of these two constraints does not make a difference. As will be shown in 2.3.3.1, however, when dealing

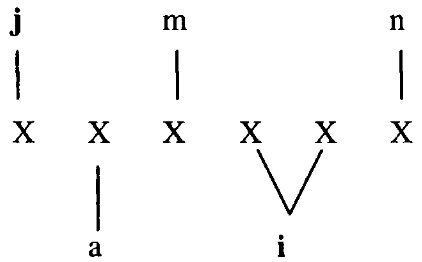
exceptions [ʔu̯iti] 'my sequins' and [bu̯iti] 'my paint' can be found. Also, a long vowel followed by a glide is ruled out, and a glide cannot be followed by the homorganic long vowel. It can, however, be followed by the homorganic short vowel, as in AA *(b)yishrab* 'he drinks' or *wuSil* 'he arrived'.

2.3.2.2 Distribution on the root and affix tiers

In the languages which have a root-and-pattern (i.e. a 'non-concatenative') morpho-phonological system, the relation between glides and vowels is complicated by the fact that vocoids can function both as root 'consonants' and as affixal 'vowels'. As was pointed out in section 1.3, Arabic has been shown to exhibit one such system (McCarthy 1981), by which the different morphemes in a word are placed on separate tiers. In particular, since Arabic roots are generally composed of non-syllabic sounds, the vowels and consonants forming a stem occupy two distinct tiers.

The distribution of glides shows that they generally behave like root consonants: they remain fixed in the various derivational forms, while the vowel pattern changes (e.g. ***wakiil*** 'agent' / ***wukalaa*** 'agents'). They must therefore be assigned to a tier other than the vocalic one, where underlying high vowels are placed (McCarthy 1981:386; Watson 1989:155). In the word *yamiin* 'right', for instance, the initial front glide and the high front vowel [i:] will belong to separate tiers (40).

(40)



These considerations would seem to call for a phonological distinction between /j/ and /i/ and /w/ and /u/. Watson (1989:165) even goes so far as to claim that glides on the root tier, like laryngeals, have consonantal features (e.g. [continuant], [sonorant], [palatal]), while underlying vowels on the vocalic tier have vocalic features. The next subsection, however, will show that things are not so simple.

---

with non-moraic O-vocoids MAX-X-IO is to be preferred, since the length rather than the weight of the

2.3.3 Glide ~ vowel alternations

The fact that high vocoids can be placed either on the root tier – where they most often function as non-syllabic sounds – or on the level 1[49] affix tier –where they most often surface as vowels – does not prevent syllabicity alternations from taking place on both tiers (and on all other affix tiers). A single high vocoid can sometimes take the form of an O-glide ([j] or [w]), of a vowel ([i] or [u]), of an N2-glide ([i̯] or [u̯]) or of a Co-glide ([j] or[w]). For instance,[50] adjectives like *abyaD* [ʔab.jaD] 'white (m.sg.)' ~ *bayDa* [be:.Da] or [bei̯.Da] 'white (f.sg.)' ~ biiD [bi:D] 'white (pl.)'; *aswad* [ʔas.wad] 'black (m.sg.)' ~ *sawda* [so:.da] or [sau̯.da] 'black (f.sg.)' ~ *suud* [su:d] 'black (pl.)' exhibit a three-way alternation, at least for those speakers who do not always monophthongize /ai̯/ and /au̯/.

In OT terms, the presence of alternations means that, as in many other languages, the NOHIATUS constraint dominates *M/I and *M/U. If the root tier glides are linked to C slots, *P/j and *P/w, which prevent glides from occupying the peak position in a syllable, must instead be outranked by SONSEQ.[51]

2.3.3.1 Alternations on the root tier

Root vocoids can occupy different positions within the syllable, according to the syllabic quality of the adjacent sounds. Alternations between O- and N2-vocoids (e.g. [w]*alad* 'boy; child' ~ *a*[u̯]*laad* 'boys; children') are fairly common,[52] and often taken for granted since both alternants are traditionally seen as glides. More noticeably, root glides are liable to alternate with the homorganic high vowels. This happens especially when the high vocoid is the last radical, i.e. when it is word-final in the unaffixed form. In words like *dalw* 'bucket' or *jidy* 'kid', the last radical surfaces as

---

relevant segments is at stake.

[49] As was said in section 1.2, level 1 affixes are those formed by vowel patterns which, combined with the root consonants, make up the Arabic word stem.

[50] The examples that will be given from now on are all from AA, if not otherwise specified.

[51] This thesis, however, generally assumes that there is no distinction between the input form of glides and that of the homorganic vowels in AA (see 2.3.5).

[52] Although monophthongization often conceals the presence of the N2 alternant (e.g. *sayf* [seef] 'sword' ~ *suyuuf* [suju:f] 'swords'; *aswad* ['aswad] 'black (m.s.)' ~ *sawda* [so:da] 'black (f.s.)').

syllabic in the isolated form or when followed by a consonant (41a,b), as an on-glide when a vowel-initial suffix is added (41c,d).

(41) (a) /dalU/ → [dɛ**lu**] 'a bucket'

(b) /ʒidI/ + /-hum/ → [ʒid**i**hum] 'their kid'

(c) /dalU/ + /-ak/ → [dɛl**w**ɛk] 'your (m.s.) bucket'

(d) /ʒidI/ + /-ak/ → [ʒid**j**ɛk] 'your (m.s.) kid'

Adjectives ending in a high vocoid in the masculine singular - which corresponds to their basic form - exhibit a clear vowel ~ glide alternation, too. As is shown in (42), while in the unaffixed adjectives such a vocoid is syllabic, when a vowel-initial suffix (in this case, the feminine singular) is added its non-syllabic form surfaces.

(42)	Hil[u]	~	Hil[w]a	'sweet, beautiful'
	ghaal[i]	~	ghaal[j]a	'expensive'
	faaD[i]	~	faaD[j]a	'empty'
	xaal[i]	~	xaal[j]a	'empty'
	etc.			

It must be pointed out, however, that there is another group of adjectives which, although ending in a high vocoid and having very similar forms to the ones in (42), do not exhibit syllabic alternation, but avoid a vowel clash by means of Glide Insertion (GI). Examples are given in (43).

(43)	ghan[i]	~	ghan[ijj]a	'rich'
	^{c}aad[i]	~	^{c}aad[ijj]a	'ordinary'
	bunn[i]	~	bunn[ijj]a	'brown'
	etc.			

This type of data definitely shows that there must be a difference in the lexical inputs of the final vocoids between the two sets of adjectives. At first sight, it may seem that the difference can only lie in the syllabicity of the two phonemes, i.e. that either (a) the final vocoids in (42) are underlyingly underspecified for syllabicity while the ones in (43) are specified as [+syllabic] (i.e. prelinked to a V position on the skeleton), or (b) the vocoids in (42) are underlyingly glides while the ones in (43) are underlyingly vowels. A closer scrutiny, however, reveals that the adjectives behaving like the ones in (43) are precisely those that exhibit a final long high vowel in SA. Therefore, it seems reasonable to assume that such a vocoid is underlyingly long in the AA adjectives, too, and is shortened only when word-final (cf. 3.3). The first segment of the geminate glide in a word like *ghani*[jj]*a* 'rich (f.sg.)' is thus the result of a diphthongization of the /I:/, the second segment the result of a further lengthening into the empty onset of the following syllable (see 3.3).

From an OT perspective, the difference between the two groups of words can be accounted for in terms of the correspondence constraints MAX-X-IO and DEP-μ-IO. The former constraint requires that skeletal slots in the input have correspondents in the output, thus militating against the shortening of the high vocoid in *ghanii* or *bunnii* (see 3.3). Clearly, this constraint must be overridden by a markedness constraint against long vocoids in word-final position, namely *FINAL-VOC: (see 3.3), which however does not apply in the case of the feminine, where the long vocoid is not in word-final position. MAX-X-IO has been preferred over the more commonly adopted MAX-μ-IO because the latter would militate against GF: since a high vocoid, when syllabified into the onset of the following syllable, loses its mora (but not its timing slot), GF would involve a violation of MAX-μ-IO (but not of MAX-X-IO). This would lead to the elimination of the winning candidate in cases such as *ghaalya*. DEP-μ-IO, on the other hand, becomes relevant when dealing with short input vocoids like those in (42), since it bans output moras which do not have input correspondents, thus militating against vocoid lengthening (see 3.3).[53] Since long vowels are universally dispreferred, a context-free markedness constraint *V: (no long vowels) can be postulated (see 2.3.1 above). In AA, as in many other languages, this will clearly be overridden by faithfulness constraints such as MAX-X-IO, since long vowels do surface. However, it becomes relevant as a trigger of diphthongization, when dealing with forms like *ghaniyya* [ɣanijje] 'rich (f.sg.)', in order to ban candidates such as

[53] Notice that, in this case, the constraint must refer to moras rather than to timing slots, since the optimal candidate *ghaniyya* [ɣanijje] 'rich (f.sg.)' has an extra timing slot – but not an extra mora – with respect to its input.

*[ɣani:je] (see (44)). NoHiatus, which has been included because of its relevance to the general issue of glide ~ vowel alternation, must be ranked rather high in AA, although its exact placement cannot be established on the basis of the specific examples that are given here. As will be shown below, the constraint Syll-Bin (syllables must be binary) is also relevant.

The evaluation of the candidates for *ghaniyya* [ɣanijje] 'rich (f.sg.)' is reported in tableau (44). Notice that Max-X-IO is determinant in the exclusion of candidate d., i.e. the one with GF.

(44)

Input: /ɣanI: + a/	Dep-μ-IO	Max-X-IO	NoHiatus	Syll-Bin	*V:
a. ☞ ɣanijja				**	
b. ɣani:a			*!	**	*
c. ɣania		*!	*	***	
d. ɣanja		*!		**	
e. ɣanija				***!	
f. ɣani:ja				**	*!

So far, no ranking can be established among the selected constraints, since candidate a. has anyway the fewest violations for each of them: the order in which the constraints have been placed in (44), and thus the choice of the fatal violations, is partially arbitrary and partially based on the evaluation of other words that are given below (see tableaux (45) - (47)).

The tableau in (45) reports the evaluation of *ghaalya* [ɣa:lja] 'expensive (f.sg.)', as an example of regular GF from a morpheme-final short high vocoid. Again, the ranking of the constraints is irrelevant[54] – and therefore chosen arbitrarily - because the optimal candidate d. has the fewest violations.

(45)

Input: /ɣa:lI + a/	DEP-μ-IO	MAX-X-IO	NOHIATUS	SYLL-BIN	*V:
a. ɣa:lijja	*!			*	*
b. ɣa:li:a	*!		*	*	**
c. ɣa:lia			*!	**	*
d. ☞ ɣa:lja				*	*
e. ɣa:lija				**!	*
f. ɣa:li:ja	*!			*	**

From the above tableaux, it is clear that the difference between the two classes of words is determined by a divergence in the input forms rather than by the action of the output constraints, thus showing that lexical storage has a seminal role in OT, too.

2.3.3.2 Alternations on the affixal tiers

As for (level 1) affixal high vowels, it is difficult to find them alternating with glides because they do not generally appear at the end of a word, and in word-medial position they are always surrounded by non-syllabic root segments.

Suffixes like [-u] 'him; his' (underlyingly /-U:/) or [-i] 'me; my' (underlyingly /-I:/) exhibit only a partial alternation. For exactly the same reason as the adjectives in (38), the final vocoid in a word like *biddi* 'I want' (underlyingly /bidd+I:/) does not undergo complete alternation with a glide before vowel-initial suffixes: the vowel being long, only its second segment is turned into a glide (cf. *bidd[ij]ak* 'I want you (s.m.)').[55] Outside the PW, however, complete alternation may apply (e.g. *bidd[j] ashuuf* 'I want to see'), because the vowel has already been shortened (cf. 2.3.4).

In OT, where the postulation of different derivational levels is not accepted, the notion that the high vocoid is only shortened in *biddi ashuuf* and not in *biddiak* can be incorporated in the constraint *FINAL-VOC:, which bans word-final long vocoids (see 2.3.3.1). This does not affect *biddiak* (see (46)), while in *biddi ashuuf* it makes the difference between the winning candidate d. and its opponent e. (see (47)).

---

[54] Notice however that, as was mentioned above, *V: must be ranked below MAX-X-IO and/ or SYLL-BIN, otherwise the candidate *[ɣa.lja] (not reported in the tableau because not relevant to this discussion) would be more harmonic than [ɣa:.lja].

[55] It must be pointed out, however, that the combination *biddu* 'he wants' + *-ak* curiously yields [biddujɛk] rather than *[bidduwɛk]. In this case, therefore, the inserted [j] can be seen as a default segment rather than as the lengthening of the preceding vocoid.

(46)

Input: /bIddI: + ak/	DEP-X-IO	*FINAL-VOC:	MAX-X-IO	NOHIATUS	SYLL-BIN[56]	*V:
a. bid.dij.jak	*!				*	
b. bid.di:.ak				*!	*	*
c. bid.di.ak			*!	*	**	
d. bid.djak			*!		*	
e. ☞bid.di.jak					**	
f. bid.di:.jak	*!				*	*

The tableau in (46) shows that SYLL-BIN must be dominated by DEP-X-IO, in order for candidate e. to be more harmonic than candidates a. and f. It must also be dominated by MAX-X-IO, so that candidate d. (i.e. the one with GF) is eliminated. Finally, NOHIATUS must dominate SYLL-BIN in order to exclude candidate b.[57]

Tableau (47) shows that *FINAL-VOC: becomes important when dealing with vocoid clashes across PW boundaries. In order for candidate d. to be selected as optimal, this constraint must dominate MAX-X-IO.

(47)

Input: /bIddI: + aʃU:f/	DEP-X-IO	*FINAL-VOC:	MAX-X-IO	NOHIATUS	SYLL-BIN	*V:
a. bid.dij.ja.ʃu:f	*!	*			*	*
b. bid.di:.a.ʃu:f		*!		*	*	**
c. bid.di.a.ʃu:f			*	*!	**	*
d. ☞bid.dja.ʃu:f			*		*	*
e. bid.di.ja.ʃu:f		*!			**	*
f. bid.di:.ja.ʃu:f	*!	*			*	**

2.3.4 Postlexical Glide Formation

As was mentioned in 2.3.3, syllabic alternations apply both within the PW (e.g. *Hil[u]* 'sweet; beautiful (m.s.)' ~ *Hil[w]iin* 'sweet; beautiful (pl.)') and across PW-boundaries (e.g. *haada ghaal[i]* 'this (m.) [is] expensive (m.s.)' ~ *haada ghaal[j] u jamiil* 'this (m.) [is] expensive (m.s.) and beautiful (m.s.)'). There are, however,

---

[56] Since word-final consonants in AA seem to be extrasyllabic (see 1.3), the last syllable in each of the candidates is monomoraic, thus violating SYLL-BIN; that is why each candidate has been assigned at least one asterisk for this constraint.

considerable differences between the process as it applies within the PW and outside the PW. As was said above (cf. 2.3.3), words which have a long high vocoid underlyingly (e.g. /bIddI:/) exhibit a vowel ~ vowel+glide alternation (e.g. *bidd[i]* 'I want' ~ *bidd[ij]ak* 'I want you (m.s.)') if the vocoid clash applies within the same PW, a vowel ~ glide alternation (e.g. *bidd[i]* 'I want' ~ *bidd[j] ashuuf* 'I want to see') across PW-boundaries. Moreover, while syllabic alternation is obligatory within the PW, it is optional – and actually applies quite rarely in a variety like AA[58] – outside this domain. Last but not least, alternation across PW-boundaries is influenced by such factors as speech rate and register, while alternation inside the PW is not. These considerations lead to the natural conclusion that, as in Italian, we are confronted with two distinct phenomena, one of which is lexical, the other postlexical. This seems to confirm that there is an initial stage of lexical syllabification in AA, too (see 1.3). In OT, however, there can only be one level of syllabification, which is given by simultaneous sifting of all candidates through the ranked set of all constraints. Therefore, the different resolution of vocoid clashes within and across words must be interpreted in terms of constraints making reference to domain boundaries (see *FINAL-VOC: above).[59]

Let us now concentrate on postlexical GF, in order to see if it is possible to find its domain of application.

As was mentioned above, in AA postlexical GF[60] is not very frequent at a normal speech rate, although at a faster rate it seems to take place more often. This is not at all surprising, if one considers the nature of the Arabic syllable. Since Arabic in general does not allow onsetless syllables, and ʔ-insertion applies to fill in empty initial onsets, one would expect the possibility of finding vowel clashes (and thus GF) across words to be completely ruled out. However in many dialects, including AA, ʔ-insertion is only obligatory in PU-initial position, so that vowel clashes may take place within PUs even across PW-boundaries. The alternation between ʔ-insertion and GF

---

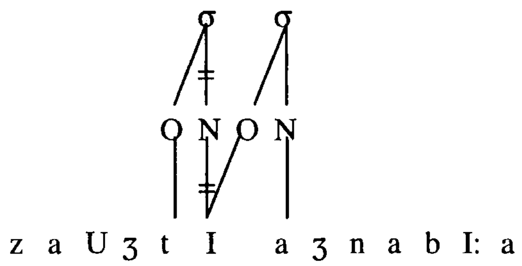
[57] The constraint *V: could also do the job, but, because of what was said in fn. 22, we have prudently ranked it below SYLL-BIN.

[58] The three AA speakers I used as controls actually produced very few examples of GF across PWs even in free speech.

[59] An OT account, however, encounters difficulties when trying to render determinant factors such as speech rate, register, etc.

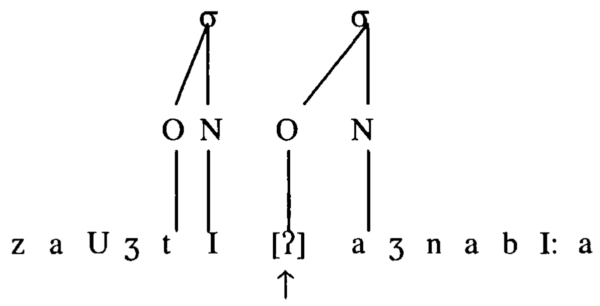
is once again dependent on the optionality of resyllabification in AA (cf. fn. 39): when resyllabification applies, a PW-final high vocoid can be resyllabified into a following PW-initial empty onset (48a); when it is blocked, GF cannot apply and a default [ʔ] goes to occupy the unfilled onset (48b).

(48) (a) Resyllabification (hence GF) applies:



'my wife is a foreigner'

(b) Resyllabification (hence GF) is blocked (ʔ-insertion applies):



'my wife is a foreigner'

Since the application of resyllabification generally depends on speech rate, so does the application of GF. It could be said that, in AA, GF across PW-boundaries applies mostly, if not only, in fast speech.[61] If GF is indeed a fast speech rule, it will be

---

[60] The term GF (Glide Formation) is adopted here because at the postlexical level, initial syllabification having already taken place, the affected vocoid is unmistakably a vowel losing its syllabicity when clashing with another vowel.

[61] Harrell (1957:68) claims that in Egyptian Arabic sequences of two successive vowels occur frequently across words. The examples he gives, however, all contain sequences of identical vowels ([di: il aʕda:d] 'these are the numbers'; [fi: itne:n] 'there are two'); in AA (and, I suspect, in many other varieties including perhaps Egyptian Arabic), though, two successive identical vowels tend to coalesce into one long vowel when a [ʔ] is not inserted (e.g. *shufti* 'you (f.s.) see' + *imtiHaanaati* 'my exams' > *shuft[i]mtiHaanaati?* 'did you (f.s.) see my exams?') .

useless to try to determine what its domain of application is, since this kind of process applies in all domains unless a pause is inserted.

In OT terms, the main difficulty lies in the fact that the two alternants are in free variation. As was mentioned in 1.5, in these cases a possible working hypothesis is that of free ranking: the candidate set is split into two (or more) subhierarchies, each exhibiting a different ranking of the relevant conflicting constraints. In this instance, the two free-ranked constraints are *M/I (no /I/ in marginal positions), which militates against GF, and DEP-C-IO (output consonants must have input correspondents), opposing Consonant Insertion. In one of the two sub-hierarchies, DEP-C-IO » *M/I, thus yielding GF (see the tableau in (49a)), in the other *M/I » DEP-C-IO, favouring ʔ-insertion over GF (see (49b)). Since the form with the hiatus never seems to surface, NOHIATUS must dominate both the other constraints.

(49) (a) First subhierarchy: NOHIATUS » DEP-C-IO » *M/I

Input: /zaUʒtI aʒnabI:a/	NOHIATUS	DEP-C-IO	*M/I
a. ☞ zo:ʒtjaʒnabijje			*
b. zo:ʒtiaʒnabijje	*!		
c. zo:ʒtiʔaʒnabijje		*!	

(b) Second subhierarchy: NOHIATUS » *M/I » DEP-C-IO

Input: /zaUʒtI aʒnabI:a/	NOHIATUS	*M/I	DEP-C-IO
a. zo:ʒtjaʒnabijje		*!	
b. zo:ʒtiaʒnabijje	*!		
c. ☞ zo:ʒtiʔaʒnabijje			*

Intonation - and thus, indirectly, sentence perspective - may also be a factor determining the application of GF, and stress assignment most probably is. Just as in Italian, GF appears to be blocked if either of the clashing vocoids bears prominence. Giving examples of stressed word-final high vocoids clashing but not undergoing GF would be meaningless, since stressed vowels are always long in such a position. The examples in (50) show instead how, even if the first clashing vocoid is unstressed, an accent on the second vocoid still blocks the process, and ʔ-insertion takes place instead.[62] By just changing the final noun from a stress-initial (50a) to a stress-final one (50b), the application of GF is made possible.

(50) (a) *kiif shuft[i] [ʔ]àasya?* 'how did you (f.s.) like Asia?' (never **shuft[j] àasya?*)

(b) *kiif shuft[j] amrìika? ~ kiif shuft[i] [ʔ]amrìika?* 'how did you (f.s.) like America?'

As in the case of Italian, GF before a stressed vowel may be prevented by high ranking of the OT constraint ANCHORING-IO(PW, L)-STRESS, by which every segment at the left periphery of the input PW that bears PW stress must have a correspondent at the left periphery of the output PW. The ranking of such a constraint with respect to NOHIATUS cannot be established on the basis of the data at hand, since both constraints are inviolate in the winning candidates. Tableau (51) illustrates the evaluation of the candidates for *shuft[i] [ʔ]àasya*, showing that candidate b., exhibiting GF, is excluded because it violates ANCHORING-IO(PW, L)-STRESS (abbreviated as ANCHOR-IO-STRESS), while candidate c. is excluded because it violates NOHIATUS. Ranking of DEP-C-IO with respect to *M/I is irrelevant in this case.

(51)

Input: /ʃUftI a:sja/	ANCHOR-IO-STRESS	NOHIATUS	DEP-C-IO	*M/I
a. ☞ [ʃufti][ʔa:sja]			*	
b.[ʃuf] [tja:sja]	*!			*
c.[ʃufti][a:sja]		*!		

Tableaux (52a) and (52b) show instead how the two renderings of *shufti amrìika* are determined by the ordering of the free-ranked constraints DEP-C-IO and *M/I: when DEP-C-IO » *M/I (52a), GF applies, when *M/I » DEP-C-IO (52b), ʔ-insertion

[62] Abumdas's (1985:165) PU-initial ʔ-insertion rule for Libyan Arabic is actually subject to the condition that the following vowel should be stressed. In AA, however, ʔ-insertion seems to apply before any vowel PU-initially and, optionally, even PU-medially.

prevails. Since no stressed PW-initial vocoid is involved, it is ANCHORING-IO(PW, L)-STRESS that is irrelevant in this case.

(52) (a)

Input: /ʃUftI amrI:ka/	ANCHOR-IO-STRESS	NOHIATUS	DEP-C-IO	*M/I
a.[ʃufti?am][ri:ka]			*!	
b. ☞ [ʃuftjam][ri:ka]				*
c.[ʃuftiam][ri:ka]		*!		

(b)

Input: /ʃUftI amrI:ka/	ANCHOR-IO-STRESS	NOHIATUS	*M/I	DEP-C-IO
a. ☞ [ʃufti?am][ri:ka]				*
b.[ʃuftjam][ri:ka]			*!	
c.[ʃuftiam][ri:ka]		*!		

In relation to SA, Benhallam (1980:57) suggests that a more satisfactory description of vowel ~ glide alternations may be couched in 'the rhythmic configurations that SA allows', i.e. in the fact that the patterns long – long – long (i.e. three heavy syllables in a row) and short – short – short (i.e. three consecutive light syllables) appear to be avoided. Similar considerations may be advanced with regard to AA where – as will be shown in more detail in sections 3.3 and 4.3 - sequences of monomoraic syllables are strongly dispreferred.

2.3.5 The phonological status of glides (and high vowels) in Arabic

The question of the phonological status of Arabic glides has been implicit throughout the previous discussion. Should we consider root vocoids as underlyingly 'consonantal' - hence more prone to occupy marginal positions in the syllable - and affixal vocoids as underlyingly 'vocalic' – and thus more liable to syllabicity - or should we regard homorganic high vocoids as exhibiting exactly the same feature matrix, and acquiring different syllabic status simply according to the quality of adjacent segments?

If we were to subscribe to a view of root vocoids as underlyingly 'consonantal' and of affixal vocoids as underlyingly 'vocalic', we should then discriminate between two separate phenomena: a process of GF from the affixal tier and a process of VF from the root tier. In fact, many scholars (e.g. Harrell 1957, Abumdas 1985, Kenstowicz 1994) regard glides as the underlying forms in this latter case.

So far, however, I have not found a definite piece of evidence that could prove that high vocoids on the affix tier have an independent phonological status from high vocoids on the root tier, apart from native speakers' intuitions (see Cantineau 1960:88) and the fact that, morphologically, root vocoids behave like consonants and not like the non-alternating (non-high) vowels.

In order to prove that root glides are independent phonemes from the homorganic high vowels, some phonologists have put forward alleged minimal pairs for some particular variety. According to Harrell (1957:28), for example, while Egyptian [i] and [j] are merely allophones of the same phoneme, /u/ and /w/ are separate phonemes with minimal pairs such as /ʕadu/ 'enemy' ~ /ʕadw/ 'running'. Contrasts such as this, however, are not so much between a vowel and a glide, as between a long vocoid and a short one: the /U/ of /ʕadU:/ 'enemy' is – at least underlyingly – long, while the short vocoid in /ʕadU/ 'running' always surfaces as a vowel unless it can be resyllabified into a following unfilled onset. In relation to SA, works as different as Cantineau's (1960), Holes's (1995) and Dickins's (1996) all confirm the notion that there are no minimal pairs within the same morphological domain, concluding that [j] and [i] and [w] and [u] respectively are allophones of the same phonemes. With regard to postlexical GF, it should also be mentioned that both root and affixal vocoids undergo exactly the same type of alternation, and that there is no sign of the former yielding non-syllabic allophones more readily than the latter.

Benhallam (1980:31-59) lays out the two different positions in order to evaluate their effectiveness in relation to SA syllable structure. According to him (ibid.: 32, following Brame 1970), the fact that glides are deleted intervocalically in lexical categories such as 'hollow verbs'[63] (e.g. /sammai̯+a/ > [samma:] 'he named',

[63] 'Hollow verbs' are verbs that are traditionally construed as having underlying roots of the shape /CajaC/ or /CawaC/. They exhibit two different patterns in the perfect tense: /CiC/ or /CuC/ respectively before consonant-initial inflectional suffixes, /Ca:C/ before vowel-initial or zero suffixes (Kenstowicz and Kisseberth 1979:415 Cantineau 1960:194), as shown in the examples below.

(iii) /ʃajal/ 'carry' →[ʃil-t] '(I) carried'/ [ʃa:l] '(he) carried'/ [ʃa:l-u] '(they) carried'
(iv) (iv) /ʃawaf/ 'see' → [ʃuf-t] '(I) saw'/ [ʃa:f] '(he) saw'/ [ʃa:f-u] '(they) saw'

/sammai+ata:/ > [sammata:] 'they both (f.) named') shows that such glides must be analyzed as underlyingly non-syllabic. This diagnostic, however, does not seem to hold, because it would be highly improbable for such vocoids to surface as vowels in such an environment. In fact, there is no need to regard the final vocoid in the stem as an underlying glide: it is simply an underspecified segment that may be realized as a vowel, as a glide or as zero according to the position in which it surfaces. The UR of a word like [samma:] will therefore be represented as /sammaI+a/.[64]

Benhallam (ibid.:53-4) further argues that vocoids which always surface as glides, even if by virtue of their pre-vocalic position, pose a challenge to the hypothesis that all glides are derived. In the analysis at hand, however, such vocoids are underlyingly underspecified for syllabicity, and are thus potentially alternating; they merely happen to always surface in a pre-vocalic environment, thus being always assigned to a syllable onset.

In conclusion, both in Arabic at large and in AA in particular, there does not seem to be any compelling evidence in favour of assigning independent phonological status to homorganic glides and high vowels, since their distribution is totally predictable on the basis of syllable structure. Neither is a distinction between 'consonantal' and 'vocalic' glides applicable to Arabic, where both root and affixal vocoids undergo syllabic alternation. Talking of GF or VF prior to initial syllabification in Arabic is thus an arbitrary stand, since all segments are underlyingly underspecified for syllabicity and homorganic high vocoids share the same feature specifications. The terms 'consonantal' and 'vocalic' tier are also misleading, and should be replaced by the terms 'root' and 'affix' tier respectively. This consideration reflects the fact that such tiers are morphological rather than phonological entities.

Intervocalic GD, however, is not limited to verbs: comparison with other derivatives from the same root reveals that words like [sama:] 'sky' (cf. [sama:wi] 'azure') or [na:b] 'canine tooth' (cf. [ʔanjab] 'canine teeth') actually display an intervocalic glide underlyingly.

[64] For a more detailed description of Glide Deletion, accompanied by a DT and an OT interpretation of the phenomenon, see 4.3.

CHAPTER 3: LENGTHENINGS AND SHORTENINGS

3.1 Long and short vowels

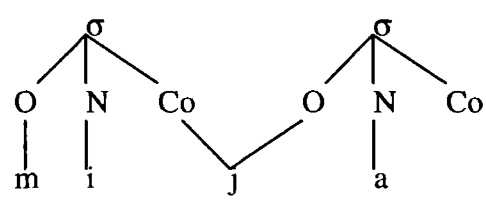
In this section, the issue of vowel length is taken into consideration. Subsection 3.1.1 gives first a brief overview of quantity in vocoids at large, including non-syllabic glides, but then focuses on vowels, looking at the representation of their length, at underlying and surface quantitative distinctions and at the factors conditioning vowel quantity, such as syllable structure and stress assignment. The second section deals with length alternations, illustrating cross-linguistically common processes such as Closed Syllable Shortening, Trisyllabic Shortening, Compensatory Lengthening and Final Lengthening. Finally, section 3.1.3 gives an OT account of vowel length distribution and alternations.

3.1.1 Vocoid length

While the length of a sound can be measured phonetically, phonological length is a relative feature, and the definition of a long vs. a short segment is given by underlying and/or surface contrasts within a given language. In phonetic terms, a (phonologically) long segment in one language can actually be far longer than a (phonologically) long segment in another language.

Both consonants and vowels can be either short or long. Therefore, although glides are intrinsically shorter than vowels, both non-syllabic and syllabic vocoids can exhibit different degrees of length. What is generally defined as a 'long glide' extends over two syllables, occupying the coda of the first and the onset of the second one, as in AA *miyya* [mij.ja] 'a hundred' (see the simplified diagram in (1)).

(1)



Such geminated glides behave exactly like double consonants, thus raising the question of whether some glides may be consonantal. As was mentioned in section 2.1, however, the syllabic placement of a segment cannot be taken as evidence of its being a vocoid or a contoid: just as contoids can be syllabic, vocoids can occupy marginal positions within the syllable. Moreover, in languages such as Arabic, double glides may alternate with (generally long) vowels (see 2.3). These considerations are in keeping with the notion that, at the segmental level, high vowels and the homorganic glides share the same feature matrix. At the syllabic level, therefore, a long vocoid may span any two sub-syllabic nodes: a long glide, as was said, extends over a Co and an O position, a long vowel over an N1 and an N2 position (see 2.1). A long vocoid may also span a sequence of a vocalic and a consonantal position (cf. AA *fii* [**fij**] 'there is (some)'), thus forming a homorganic diphthong (see 2.3).
This chapter, however, concentrates especially on long vowels, which are the most prototypical long vocoids, and on their alternation with short vowels.

3.1.1.1 Representation of long vowels

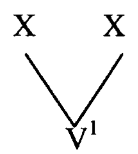
Long vowels have been represented in various ways, and particularly as a sequence of two identical short vowels versus 'a single segment marked by a feature of length' (V: or [+long]) (Donegan 1985:98). Both Kenstowicz (1970) and Pyle (1970) present evidence, in the shape of prosodic and segmental processes, for both types of representation (in Donegan 1985:98-100).
Harrell (1957:52) argues that a 'length phoneme' notation (i.e. V:) is to be preferred on the basis of the fact that it offers the following advantages:

1. It reduces the number of symbols and makes transcription easier.
2. It makes possible a simpler and more consistent description of stress and morphophonological alternation.

As was pointed out above and in 2.1, the second element of a long vowel can be regarded as non-syllabic. This, however, does not prove that long vowels are to be considered as sequences of two short vocoids, since short vowels, too, may be treated as formed by two separate elements, i.e. they may undergo diphthongization (cf. short diphthongs in languages like Icelandic and Old English). Moreover, diphthongs themselves can be short or long. Donegan (ibid.: 104) therefore concludes that 'it is not possible to do without a prosodic specification of length'.

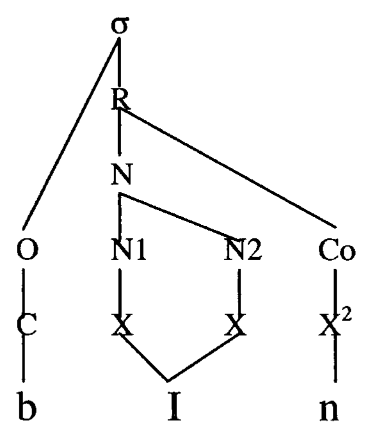
An autosegmental representation, in fact, can solve the problem in a more elegant way. In autosegmental terms, a long vowel can be regarded as a single vocoid linked to two consecutive timing slots - or 'syllabic terminal nodes' (Prince 1984) - on the skeleton, as in (2).

(2)



The use of the distinctive feature [long] is therefore to be rejected in favour of a representation in which vowel – and, more generally, segment – length is located on the skeletal tier. Moreover, having an independent tier for segmental length can allow for a better characterization of its connection with syllable structure, as in the following representation of *bean* [bi:n].

(3)



The notion lying at the basis of such a representation is that of an OCP by which two consecutive identical segments are banned, and long segments must be regarded as

[1] The symbol 'V' here stands for the feature specifications of any vowel, rather than for a vocalic slot on the skeleton.

[2] English /n/ has been assigned an X slot on the skeleton because it may alternate between a non-syllabic and a syllabic position (e.g. *botton* [bʌtn̩]).

single segments attached to two timing slots (see for example McCarthy 1981, Prince 1984).

Prince (1984:240) supports what he calls a 'sequential' representation of long vowels - i.e. one in which length is reported on the skeletal tier - on the basis of the fact that the presence of such segments in a language does not necessarily imply the presence of all diphthongs combining any two short vowels. In Finnish, as in many other languages, only high vocoids may form the second element of a diphthong, while any vowel may be long.[3] In linear phonology, this would require 'a clumsy disjunction': in tautosyllabic V_1V_2 sequences, either V_2 is [+high] or $V_1 = V_2$. An autosegmental representation, on the other hand, allows one to maintain quantitative distinctions at the skeletal level, while disallowing (at least certain) vocoid sequences at the segmental level. The disjunction is thus avoided, because the constraint on diphthongs belongs to the segmental tier, while gemination is defined on the skeleton.

3.1.1.1.1 The mora

As was mentioned in 1.3, vowel quantity, along with syllable weight, can be expressed in terms of moras. A short vowel can thus be regarded as associated to one mora, a long vowel to two moras. Within a theoretical framework including a skeletal tier, however, since long and short vowels are already distinguished by the number of timing slots they are linked to, the role of the mora in this sense becomes redundant. Moreover, the number of moras cannot discriminate between a long consonant spanning a coda and an onset and a short coda consonant, since both of them are monomoraic. Therefore, this thesis does not refer to the mora as a unit of length, let alone as a prosodic constituent, but rather as a unit of syllabic weight. Length is instead represented on the CVX-tier. However, there are processes and requirements that need to refer to the weight, rather than to the length, of segments. The constraint on foot binarity (see 1.4, 1.5), for instance, crucially refers to the notion of mora, in that it states a preference for disyllabic or bimoraic feet. Skeletal length, on the other hand, is clearly irrelevant to this constraint: a foot such as $[\text{stru}]_F$ is degenerate (i.e. unary) in spite of the fact that it spans four timing slots.

[3] Indeed, in a language with contrastive vowel quantity diphthongs may be altogether absent.

In any case, both moras and skeletal positions may be used to represent and account for vowel length, since in most languages short vowels may be seen as monomoraic, long vowels as bimoraic.

3.1.1.2 Underlying vs. surface length

As was mentioned above, vowel length can be either phonemic or allophonic. Among the languages which typically do not exhibit contrastive length are most Austronesian and Sino-Tibetan languages (De Chene 1985:11-13), and Romance languages such as Italian (see 3.2) and French[4].

In other languages vowel length is contrastive, or lexical, and is thus to some degree independent of the environment. Most of these distinguish between only two degrees of phonological length, i.e. only between short and long vowels (Donegan 1985:97). For a few languages, such as Estonian, Mixe and Yavapai, it has been claimed that there are three contrasting vowel lengths (Ladefoged and Maddieson 1996:320). In the case of Estonian, however, Lehiste (1970, in Ladefoged and Maddieson 1996) argues that such a distinction is partly dependent on word structure, and cannot therefore be considered purely lexical. In Yavapai, and especially in Mixe, on the other hand, the three-way length distinction appears to be independent of morpheme structure or word patterns (Ladefoged and Maddieson 1996:320).

Apparent cases of four contrasting lengths can instead be explained on the basis of the distinction between tautosyllabicity and heterosyllabicity (ibid.:320-1). In Kikamba, for instance, both long and short identical consecutive vowels may be separated by a syllable boundary, so as to yield the types V, V:, V.V and V.V: (V:.V and V:.V: are both absent), where V.V is considerably longer than V: (Roberts-Kohno 1994, ibid.).

The status of vowel length in a given language can of course change over time. De Chene (1985) maintains that the only possible way in which a language which does not distinguish between different vocalic lengths can acquire such a distinction diachronically is by means of coalescence of vowel sequences, often after the loss of an intervocalic consonant. However, it can be argued that in some languages (German, English, with further substitutions and diphthongizations) length has become contrastive as a consequence of the degemination of following consonants, where

[4] A quantitative distinction has been preserved in certain regional varieties of French (e.g. Ardennes, Bourgogne, Berry, Loire, Normandy) and Italian dialects (see 3.2), but not in the standard forms.

temporal compensation (Compensatory Lengthening) does not take place (see below). For instance, the short [i] in German *riffeln* [Rifəln] 'to card' has come to contrast with the long one in *riefeln* [Ri:fəln] 'to groove, plough'.

3.1.1.3 Factors conditioning vowel length

3.1.1.3.1 Intrinsic duration

Donegan (1985:92) points out that, even in languages exhibiting contrastive length, the duration of a segment is determined by factors other than its lexical collocation, such as the environments in which it surfaces or its intrinsic phonetic properties. While the first two factors are definitely fundamental in a phonological perspective - in that they are strictly connected to the distinction between surface and underlying length – the latter seems to be to a large extent irrelevant to our purposes here. However, Donegan (ibid.:92-4) maintains that intrinsic duration, though never contrastive, can actually influence the distribution of phonemes[5] and the application of phonological processes which are sensitive to length.

She (p. 92) reports that intrinsic duration is – universally, it seems - related to other phonetic features, such as vowel height, tenseness, timbre and colour[6]: lower vowels are longer than the corresponding higher vowels, tense vowels are longer than the corresponding lax vowels, chromatic vowels are longer than the corresponding achromatic vowels. Donegan (ibid.:53-4) further argues that vowel length is also related to sonority, i.e. 'the more sonorant a vowel is, the more susceptible it is to lengthening, and conversely, the less sonorant it is, the more susceptible to shortening (or loss)'. With reference to chapter 2, one may wish to add 'or to desyllabification'. These are all reduction or fortition processes, to do with the 'sustainability' or 'continuance' of vowels (ibid.:53). Donegan (p. 54) points out, however, that 'the relation of sonority to sustainability may be mediated by intrinsic length, since more sonorant vowels are intrinsically longer than less sonorant vowels'. Also, more sustainable vowels promote continuance in adjacent consonants (ibid.).

---

[5] For example, in many languages features like [long] and [high], or [short] and [low], or [long] and [lax], or [short] and [tense] do not co-occur (ibid.:93).

[6] By 'colour', Donegan means the degree of palatality or labiality exhibited by a vowel. For instance, a vowel like [æ] is 'chromatic' – i.e. it has a palatal quality – while [ɑ] is not.

3.1.1.3.2 Relation with stress distribution

As was pointed out in 1.4, accent placement is tightly related to vowel length, both because long vowels, by making up heavy syllables, attract stress and because stress assignment triggers vowel lengthening. Quite obviously, while all stressed vowels are phonetically lengthened, vowel length can only determine stress placement in those languages in which length is contrastive.

Anderson (1984), going back to a proposal first laid out by Jakobson (1928), maintains that no language can have both 'free quantity' – i.e. contrastive length – and 'free dynamic stress' – i.e. contrastive stress accent. This, according to the author, is 'a direct consequence of the nature of these properties themselves' (ibid.:105), provided they are structurally motivated on the basis of a metrical approach. According to such an approach

> A language with free quantity is one that displays in its surface forms a lexically specified prominence distinction between internally complex vowels and internally simple ones. A language with free stress is one in which this same distinction is neutralised [...]. Manifestly, the two situations are not compatible.

These considerations, however, seem to be far too categorical. In a language like English, for instance, vowel quantity depends to a great extent on stress (Myers 1987:486), even though, as was mentioned above, it can actually be contrastive (cf. *bean* [biːn] vs. *bin* [bɪn])[7].

Myers (1987:495) mentions a stress sensitive rule of resyllabification, by which a sequence V.CV is resyllabified as VC.V when the first vowel is stressed and the second is unstressed (e.g. *vánity* ['væ.ni.ti] > ['væn.i.ti]). Such a process, which applies in languages such as English, Danish and Proto-Germanic, also relates to the tendency for stressed syllables to be longer than, and even attract material from, unstressed syllables.

Many languages, among which Danish, Yupik, Biblical Hebrew and Italian (cf. 3.2), exhibit a constraint by which a stressed rhyme always branches (Myers 508), i.e. which bans monomoraic stressed syllables *(C)V́. This is clearly a corollary of the universal constraint against monomoraic syllables that was mentioned in section 1.3,

[7] The contrast in tenseness between [iː] and [ɪ] can be seen as a consequence of that in length.

and is often referred to as 'Bimoraic Enforcement' or 'Strong Rhyme Condition' (see for example Chierchia 1986:22). The tendency to avoid monomoraic syllables is in its turn related to Broselow's (1992) *Bimoraicity Constraint*, by which all syllables are maximally and optimally bimoraic (see 1.3). These requirements are enforced more readily in stressed than in unstressed syllables because segmental length and syllabic weight are among the clearest correlates of phonetic prominence.

3.1.1.3.3 Relation with syllable structure

As is apparent from what has been said, syllable structure is a seminal factor in determining the length of a vocoid. In languages such as Swedish, for example, consonants and vowels may both be long, but a long vowel and a long consonant cannot co-occur within a rhyme, i.e. the rhymes with long segments can take the form V:C or VCC, but not V:CC (Carr 1993:210). This goes to show that there is a restriction on the number of timing slots in the rhyme, so that the presence of a long vowel precludes the occurrence of a long consonant in the same rhyme, and vice versa (ibid.:211). In terms of moras, in Swedish trimoraic rhymes (or syllables) are allowed, while rhymes (syllables) comprising four or more moras are banned.

In English, on the other hand, trimoraic syllables are generally avoided, too. In this language, long vowels appear only in open syllables word-internally, and may be followed by a tautosyllabic consonant only when this is word-final, as in *tape* or *seem*. In the latter cases, the final consonant can be construed as extrametrical, thus maintaining the generalization that trimoraic syllables are disallowed (Myers 1987:488; see also Broselow 1992:10, Abu-Mansour 1992:53, Watson 1999a:503).[8]

These considerations directly relate to the above-mentioned universal preference for bimoraic syllables, i.e. to the Bimoraicity Constraint. In particular, they point to a ban against trimoraic syllables which is cross-linguistically stronger than that against monomoraic syllables.

A phonetic reflex of this prohibition is the fact that, universally, vowels in closed syllables have a shorter duration than vowels in open syllables (Maddieson 1984). Donegan (1985) ascribes this phenomenon to a general tendency to 'temporal compensation'. This type of temporal compensation, however, is far more evident in

[8] Since, however, examples like *kind* [kʌɪ̯nd] do exhibit trimoraic syllables (even if the last consonant is extrametrical, the rime is still made up of three segments), the ban cannot be absolute.

syllable-timed languages, where syllables all tend to be of equal length (see 3.2). In stress-timed languages, on the other hand, another type of temporal compensation takes place, by which the accented vowel of a monosyllabic word will, other things being equal, be longer than that of a disyllabic or trisyllabic word. Beckman and Edwards (1990), in fact, refer to this durational effect as to 'stress-timed shortening'. They (ibid.) suggest, on the grounds that it does not refer to prosodic boundaries but only to prominence peaks, that this phenomenon should be seen as a mere rhythmic effect rather than as a prosodic operation. It must be pointed out, however, that in many languages, and particularly in stress-timed languages, both 'syllable isochrony' – by which all stressed syllables are long and all unstressed syllables are short – and 'stress-group isochrony'[9] – by which stressed vowels in polysyllabic groups are shorter than those in monosyllabic groups – take place (Donegan 1985:96).

According to Donegan (1985:96-7), temporal compensation does not mark lexical distinctions, but can condition (i.e. feed) phonological processes such as diphthongization (see 2.2), palatalization, labialization.

3.1.1.3.4 Influence of adjacent sounds

Another factor that may determine the length of a vowel is the quality of adjacent segments. In most varieties of English, for instance, vowels are phonetically lengthened when followed by a voiced consonant. A word like *bead* [bi:d] thus exhibits a longer vowel than one like *beat* [biˑt]. While this appears to be a mere phonetic effect, in Lowland Scots a process of lengthening before voiced continuants applies in a more systematic way. In this variety, as can be seen in (4), vowel length is allophonic, and long vowels only surface root-finally or before a tautosyllabic voiced continuant (Carr 1993:92-6).

(4)	*long vowels*	*short vowels*
	heave [hi:v]	beef [bif]
	swathe [swe:ð]	faith [feθ]
	boor [bʉ:ɹ]	loot [lʉt]
	hoe [ho:]	oath [oθ]

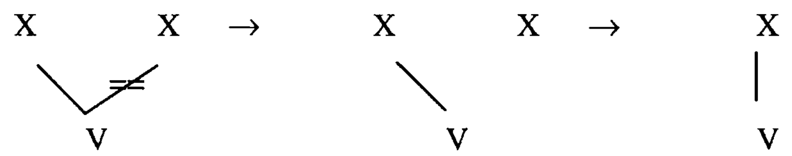
(Adapted from Carr 1993:93-4)

[9] Donegan's 'stress-group' seems to correspond to the prosodic foot.

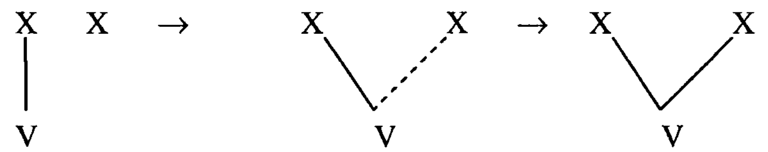
3.1.2 Length alternations

Long and short vowels alternate both in languages that exhibit an underlying distinction and in languages that do not. The process by which length alternation takes place can be characterized either as lengthening – and this is always the case in languages with no underlying length contrasts – or as shortening, according to what the UR is deemed to be. While lengthenings are strengthening, or fortition, processes, shortenings are weakening, or reduction, processes. In autosegmental terms, the latter are regarded as deletions of a timing slot on the skeletal tier (5a), the former as insertions of a timing slot (5b).

(5) (a) Vowel Shortening:



(b) Vowel Lengthening:



Although diachronically lengthening and shortening present themselves as two clearly distinct (indeed opposite) processes, synchronic alternations are often difficult to classify as due to one or the other phenomenon.

The following subsections look at a few cross-linguistically common phenomena involving vowel length alternations. As was anticipated above, alternations are typically triggered by syllable and foot structure constraints, and by stress shifting.

3.1.2.1 Closed Syllable Shortening and Open Syllable Lengthening

As was explained in 3.1.1.3.3, vowel length is often connected to a cross-linguistically common preference for bimoraic syllables. Since short vowels and coda consonants are typically assigned one mora, while long vowels span two moras, an optimal bimoraic rhyme will comprise either a short vowel and a coda consonant or a long vowel on its own. A V:C rhyme is trimoraic, and thus violates the Bimoraicity Constraint. By the same token, an open syllable whose rhyme is only made up of a short vowel is also dispreferred, being monomoraic. In the first case, in order to achieve the well-formed bimoraic syllable, the vowel may be shortened. This process is often referred to as *Closed Syllable Shortening* (CSS). In the second case, and for the same reason, the vowel may be lengthened, thus undergoing what we may call *Open Syllable Lengthening* (OSL). These processes being two sides of the same coin, knowing whether an alternation is the result of CSS or OSL may be irrelevant or even impossible. In certain cases, however, it is clear that one of the two processes is to be selected because the environment for its application is met at some point in the derivation, for example as a consequence of affixation. In English, for instance, CSS applies 'when in the course of the root-level derivation a CVV syllable comes to be followed by a C [consonant] which cannot be incorporated into any following syllable', and cannot be extrametrical because of its non-peripheral position, as in *deceive* [dɪsi:v] ~ *decep-tive* [dɪsɛptɪv], or *seduce* [sɪdju:s] ~ *seduc-tion* [sɪdʌkʃən] (Myers 1987:488-9).

CSS is a widespread phenomenon, which can be found in such different languages as English, Hausa, Turkish and Yawelmani (Myers 1987:507). As was mentioned above, there is a universal tendency for vowels in closed syllables to be shorter than vowels in open syllables.

It can be formulated as in (6) or, in autosegmental terms, as in (7).

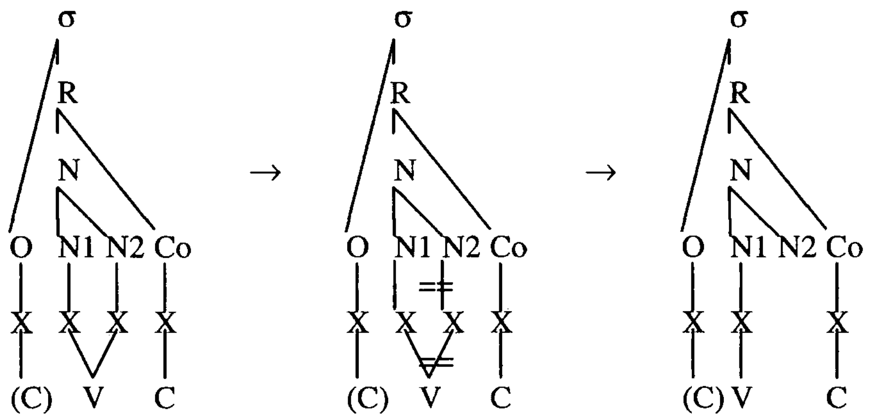
(6)

Closed Syllable Shortening (I):

$V: \rightarrow V \,/\, \underline{\quad\quad} C]_\sigma$

(7)

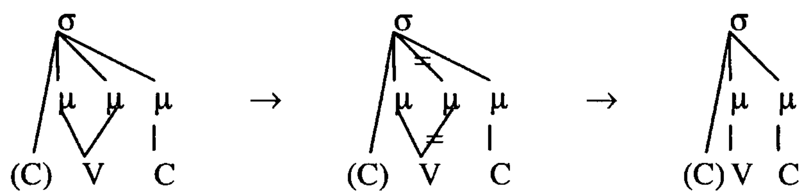
Closed Syllable Shortening (II):



It must be pointed out, however, that what the process refers to here is not so much subsyllabic constituents, but syllable weight. The representation in (7) fails to show that the N2 node is delinked whenever the Co node is linked to segmental material, precisely because it fails to show any connection between these two nodes apart from the indirect one through the R. The skeleton as represented in (7), on the other hand, though indicating that one of the X-slots is deleted in the process, does not rule out the possibility that deletion of the onset material could obtain the same result as deletion of N1 material, since the onset node is assigned one X just like the other subsyllabic nodes. In a moraic representation, on the contrary, only N1, N2 and Co segments are assigned moras, while the onset is left weightless (8a). Since each syllable optimally contains two moras, any extra mora will be deleted together with the attached segmental material (8b, c).

(8) *Closed Syllable Shortening (III):*

(a) (b) (c)



This is evidence that a moraic tier must be included in syllabic representations, along with CVX-tier and sub-syllabic nodes.

Myers (ibid.:511), however, draws the reader's attention to the fact that a formulation of CSS as in (7) – as well as one as in (8) - presupposes a stage in the derivation where long vowels occur in closed syllables, thus positing 'an otherwise unmotivated syllable type, which is then subject to absolute neutralization'. In fact, it is sufficient to encode the constraint against CVVC syllables in a syllable template such as that in (9).

(9)

$$\sigma \rightarrow \mu(\mu)$$

This template is essentially saying that, in the relevant languages, syllables are minimally monomoraic and maximally bimoraic.

3.1.2.2 Trisyllabic Shortening

Vowel length is not only conditioned by syllable structure, but also by foot structure. In particular, in many languages the stressed vowel in trisyllabic feet[10] tends to shorten. This phenomenon is often referred to as Trisyllabic Shortening (TS), especially un English phonology (see for example Chomsky and Halle 1968).[11] In this language, TS accounts for alternations such as *profane* [prɒ'fe̯ɪn] ~ *profanity* [prɒ'fænɪti], *type* ['ta̯ɪp] ~ *typify* ['tɪpɪfa̯ɪ] or *omen* ['əʊ̯men] ~ *ominous* ['ɒmɪnəs]. Notice that the shortening cannot be triggered by syllable structure constraints, since it applies to vowels in open syllables; neither is it to do with stress assignment, since the affected vowel is stressed in both alternants within each pair. The (diphthongized) vowels are shortened when followed by two syllables within the same foot.

TS can thus be formulated as in (10).

---

[10] Remember that trisyllabic feet are allowed in this framework.

[11] Myers (1987:494-507) argues that TS is to be construed as an instance of CSS on the grounds that, in alternations such as *profane ~ profanity, type ~ typify* or *omen ~ ominous*, the relevant vowels are in a stressed syllable followed by an unstressed one. In English, as was said above, a process of resyllabification seems to create a sequence VC.V out of V.CV, provided the first vowel is stressed and the second is unstressed, thus closing the first of the two syllables and consequently triggering CSS (ibid.:195). It may be argued, however, that a word like *omen* should also undergo TS, since it exhibits a stressed syllable followed by an unstressed one.

(10) Trisyllabic Shortening:

V: → V/ [____ σσ]F

It has been pointed out (see Kiparsky 1982a, b) that not only is English TS morphologically conditioned (it does not apply in underived forms such as *nightingale* [naitiŋgeil]), but it also has exceptions, such as *obese* [əʊbi:s] ~ *obesity* [əʊbi:sɪti] (not *[əʊbɛsɪti]). On these grounds, it can be considered a lexical, rather than a post-lexical, process.

3.1.2.3 Compensatory Lengthening

A typical phenomenon of temporal compensation is Compensatory Lengthening (CL), i.e. a process by which the deletion of a post-vocalic consonant (typically, a syllable-final one) is accompanied by the lengthening of the preceding vowel. In Komi, for instance, the /l/ in verb stems such as *kïl* 'hear' and *sulal* 'stand' is deleted in pre-consonantal position; concomitantly, the preceding vowel is lengthened, as in the infinitive forms in (11).

(11)

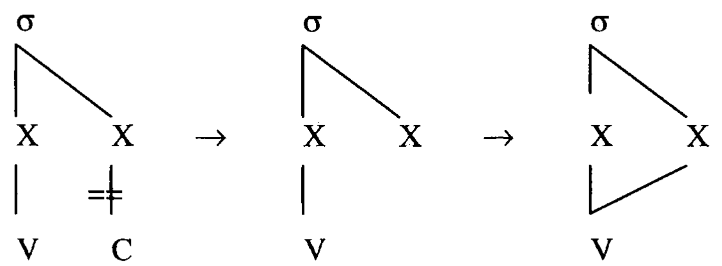
kïl	+	nï	→	kï:-nï	'to hear'
sulal	+	nï	→	sulo:-nï	'to stand'

(Kenstowicz 1994:295, from Harms 1968)

The vowel lengthening is clearly making up for the loss of the consonant, so that one can assume that the latter segment has left behind an empty position, which is then filled by the vowel.

Prince (1975, in Prince 1984) first introduced the idea that a segment (in this case a consonant) may be deleted only at the segmental level but not at the quantitative level, so as to leave an empty timing slot behind which is subsequently occupied by other material (in this case, the second element of a long vowel) by means of segmental spreading (cf. (12)).

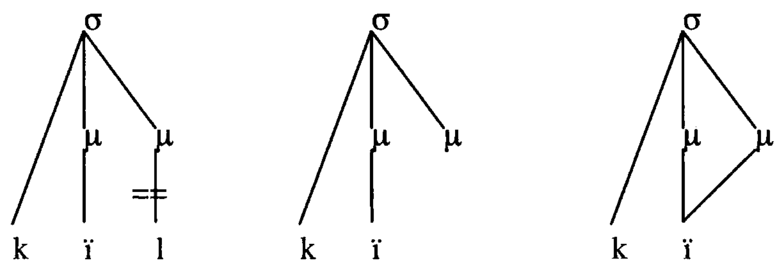
(12)



(adapted from Prince 1984)

Since the lost consonant typically belongs to the same rhyme as the lengthened vowel, CL is clearly triggered by syllable structure requirements similar to those triggering CSS and OSL: by losing a coda consonant, the syllable would become monomoraic, thus contravening the Bimoraicity Constraint, if the preceding vowel did not intervene to rescue the floating mora. In fact, the deletion of an onset consonant, which is not moraic, does not generally trigger CL. By making exclusive use of skeletal slots, Prince misses this generalization, which can only be captured by a moraic representation as the one in (13).

(13)



According to De Chene (1985:160; 189-195), for CL to take place in a given language, such a language must exhibit an independently motivated length contrast. De Chene (ibid.:198-202) also points out that certain processes, among which degemination, never seem to lead to CL.

3.1.2.4 Final lengthening

Duration is not only relevant inside the PW, but also in larger prosodic domains. Certain phrasal constituents, for instance, undergo a process of Final Lengthening (FL), which generally affects the string from the last stressed vowel onwards, and in particular the last stressed vowel itself. Although FL is a widespread phenomenon, its domain of application appears to vary from one language to the other. In English, for instance, FL seems to apply in the PPh (NV:178).

Beckman and Edwards (1990), however, claim that there are two separate processes of FL in English, which they call 'phrase-final lengthening' and 'word-final lengthening'.

> Phrase final lengthening occurs at intonational-phrase boundaries, and is a large effect that is largely consistent across speakers and rates. Word-final lengthening occurs at some other constituents' boundaries, and is a much smaller effect that is not consistently evident across speakers and rates.
>
> (Beckman and Edwards 1990:176)

Because of its high variability, it is difficult to establish what the domain of application of this latter phenomenon is. Moreover, it should be kept in mind that, in languages such as English, or Italian, or Arabic, all constituents above the PW are liable to undergo some degree of FL, since it is the last immediate constituent that tends to be stressed, provided other factors such as emphasis do not intervene.

3.1.2.5 Vowel length and syllable count: the case of Yidiɲ

An interesting case of lengthening can be found in Yidiɲ, a language spoken in Central Australia (cf. Dixon 1977, in NV 34-5, 134-5). In Yidiɲ vowel length, which is contrastive, is conditioned by syllable count. For example, there is a process that lengthens the penultimate vowel of any non-derived word with an odd number of syllables (cf. (14a) and (14b)). In derived words, if a monosyllabic suffix is affixed to a root with an even number of syllables Penultimate Lengthening (PL) regularly applies (cf. (14c)). On the other hand, if a disyllabic suffix is affixed to a root with an odd number of syllables it is the penultimate vowel of the root, and not that of the suffix, that is lengthened (cf. (14d)). Finally, if a word contains a sequence of a disyllabic plus a monosyllabic suffix, such a sequence behaves like a non-derived word as regards PL (cf. (14e)).

(14)

a. maɖi:ndaŋ 'walk up'

b. gumbiraŋa 'pick up'

c. gali:-na 'go (purp.)'

d. maḏi:nda-ŋaliŋ 'walk up (pres.)'

e. guma:ri-daga:-ɲu 'to have become red'

3.1.3 Vowel length in OT

As was mentioned in 2.3, it has been observed that short vowels are universally preferred to long vowels; hence, they can be construed as unmarked and a general constraint *V: (no long vowels) can be postulated.

3.1.3.1 Length and syllable structure

Since the length of a vowel is very often related to the overall moraic structure of the syllable, the preference for short vowels is also partially reflected in the constraint *3μ (no trimoraic syllables). This bans long vowels in closed syllables, because they would contribute two extra moras to the already moraic coda consonant(s). On the other hand, monomoraic syllables are also cross-linguistically dispreferred, as stated in the constraint $*\sigma_{\mu}$, which can be a trigger to vowel lengthening. Both *3μ and $*\sigma_{\mu}$ can be subsumed under the more general constraint SYL-BIN, requiring syllables to be bimoraic and thus reflecting Broselow's (1992) Bimoraicity Constraint. However, since the ban against trimoraic syllables is often stronger than that against monomoraic syllables, it is preferable to keep *3μ and $*\sigma_{\mu}$ as separate constraints.

The constraint DEP-μ-IO (or DEP-X-IO, see 2.3) militates against vowel lengthening, stating that 'output moras [or timing slots] have input correspondents'. More in general, both vowel lengthening and vowel shortening are disallowed by the following identity constraint:

WT-IDENT-IO

If $\alpha \in$ Domain(*f*),

if α is monomoraic, then $f(\alpha)$ is monomoraic. ('no lengthening')

if α is bimoraic, then $f(\alpha)$ is bimoraic. ('no shortening')

(McCarthy 1995a)

The fact that, in many languages, coda consonants are assigned a mora (while onset consonants are not) is captured by the constraint WEIGHT-BY-POSITION ('coda consonants are moraic'). However, its effects may be neutralized in word-final position if it is dominated by the constraint *FINAL-C-μ, by which the final consonant is weightless, a kind of NONFINALITY constraint that describes the extrametricality of a word-final consonant (see 1.5).

3.1.3.2 Length and foot structure

Lengthenings and shortenings often apply to satisfy foot structure requirements, such as the anti-degenerate-foot constraint FT-BIN (feet are binary under moraic or syllabic analysis). In languages in which FT-BIN » DEP-μ-IO, lengthening may apply in order to avoid monomoraic feet. Another foot well-formedness constraint is UNEVEN-IAMB ((LH) ≻ (H), (LL)), which is derived from the notion that the cross-linguistically unmarked iambic foot is formed by a light plus a heavy syllable (LH). When UNEVEN-IAMB » DEP-μ-IO, the vowel of the second syllable may be lengthened (iambic lengthening).[12] On the other hand, the unmarked trochaic foot is 'even', i.e. formed by two light syllables (LL) or by two moras (H). Therefore, a long vowel may undergo shortening when the resulting light syllable is footed with a following light syllable to form an even trochee, as in English *cone* [kəʊn] ~ *conic* [kɒ.nik].[13]

Kager (1993, 1995, in Kager 1999:173-4) suggests that both iambic lengthening and trochaic shortening are due to high ranking of a constraint called RH-CONTOUR, by which 'a foot must end in a strong-weak contour at the moraic level'.[14] The cross-linguistically unmarked feet are in fact (H), (LL) and (LH) which, as is shown in (15), all end in a strong-weak moraic contour.

---

[12] Kager (1999:150) points out that a rule-based DT analysis – apart from being more abstract - is unable to give such a unified account of the formation of the canonical iamb (LH), which must be characterized by means of two seemingly unrelated rules, one assigning foot structure (metrical parsing) and the other lengthening vowels in strong open syllables (quantitative adjustment). OT, on the other hand, can provide a unified account of the phenomenon by virtue of its parallel evaluation of prosodic well-formedness and input faithfulness.

[13] Notice that FT-BIN implies EVEN-TROCHEE, but not UNEVEN-IAMB.

[14] According to Kager, this constraint is grounded in sonority sequencing, since the vocalic peak, which is more prominent than the following segments, is situated as early as possible within the nucleus).

(15) (a)	s	w	(b)	s	w	(c)	w	s	w
	[μ	μ]		[μ]	[μ]		[μ]	[μ	μ]
	(H)			**(L**	L)		(L	**H)**	

In languages in which RH-CONTOUR is ranked low, the anti-rhythmical feet (HL) and (LL) are preserved, which gives the impression of quantity-insensitive stress (Kager 1999:174).

A related trigger of lengthening is satisfaction of the minimal word, which generally corresponds to a foot. This conclusion derives by transitivity from the two requirements that the grammatical words must have prosody (GW=PW) and that a prosodic word must have a foot to function as its head. Subminimal lengthening is triggered by the following constraint ranking: GW=PW, FT-BIN » DEP-μ-IO.

Kager (1999:175) points out that OT captures the notion that quantity-sensitivity is not an on-off property of languages to be translated into a parameter, but it is gradient. Constraint re-rankings produce the various degrees of quantity-sensitivity that are found in natural languages. He also points out that quantity-sensitivity is never completely 'switched off', and that in all languages there is always some aspect in which quantity distinctions matter; this, again, is very well captured by OT's violable but universal constraints.

3.1.3.3 Length and stress

As was mentioned above, and in 1.4, stress tends to fall on heavy syllables, and hence on long vowels, and at the same time stressed vowels tend to lengthen, while vowels in unstressed syllables tend to shorten. Therefore, another constraint that may trigger vowel lengthening is the Weight-to-Stress-Principle (WSP), which enforces the link between stress and vowel length by requiring that heavy syllables be stressed (see 1.5). However, if *CLASH » WSP, some heavy syllables may lose their prominence. As was mentioned in 1.5, the mirror-image constraint is SWP (Stress-to-Weight), by which stressed syllables are heavy. If both SWP and *3μ are ranked high in a given language, their combined effect is to make all stressed syllables bimoraic. This may be achieved by vowel lengthening, if DEP-μ-IO is dominated by SWP and *3μ.

In languages which exhibit allophonic vowel length, i.e. in which the length of a vowel entirely depends on syllable structure and stress assignment, WT-IDENT-IO must be dominated by constraints such as SWP, *3μ and WEIGHT-BY-POSITION. With a similar ranking, underlying vowel length is irrelevant to the outcome of the evaluation process: regardless of whether the input vowel is long or short, the output vowel will be long if and only if it is in a stressed open syllable (cf. Italian in 3.2).

3.1.3.4 Split and contraction

A long vowel, rather than being shortened, may be broken up into two short vowels belonging to different syllables, as in the Tongan example in (16). This may be called a *split*.

(16) /hu:/ → [hu:] 'to go in'

/hu:-fi/ → [hu.u.fi] 'to open officially'

(from Kager 1999:176, from other authors)

OT accounts for this preference by ranking MAX-μ-IO - requiring all input moras to have correspondents in the output - above ONSET. The opposite process, by which two identical adjacent vowels are combined into a long vowel, is called *contraction*, and can be rendered in OT terms by ranking ONSET (or NOHIATUS) and MAX-μ-IO above MAX-σ-IO.

3.1.3.5 Compensatory Lengthening and the problem of opacity

CL is a clear case of opacity, i.e. of a phenomenon involving alternations that can only be accounted for in terms of generalizations that are not apparent in the surface forms. In particular, the lengthening of the vowel can only take place as a result of the deletion of the following consonant, and thus presupposes an intermediate stage in the derivation which is not directly reflected on the surface.

As was mentioned in 1.5, opacity is one of the thorniest problems for OT, which does not allow for intermediate levels of representation. Apart from far-fetched and ultimately unsatisfactory proposals such as McCarthy's (1998) Sympathy – advancing the idea of faithfulness to other (arbitrarily chosen) candidates - the only viable solution to this problem appears to be the actual introduction of intermediate levels of representation into OT, as proposed by Kiparsky (2000). This suggestion, however, clearly weakens the theory.

3.2 Long and short vowels in Italian

This section gives a general account of vowel length in Italian and of the processes involving it. The first subsection shows that long vowels are not part of the phonological inventory of this language, and that their distribution is totally predictable on the basis of syllable structure, stress assignment and morphological considerations. The second subsection analyzes the length alternations that take place in Italian and the factors underlying them. An alternative OT account is given for each of the issues at stake.

3.2.1 Non-contrastive length

In Italian, each of the seven vocalic phonemes can surface as short or long, according to the context in which they are found (see (17)). Thus, sonority does not seem to play any role in relation to vowel length, although, as was anticipated in 1.1, it could be argued that part of the reason why glides are only formed from high vowels is that such vowels, being less sonorous, are intrinsically shorter than non-high vowels.

(17)

	Long vowels	*Short vowels*
a.	c[a:]lo 'decrease'	c[a]láre 'to decrease'
b.	p[ɔ:]ro 'pore'	p[ɔ]rro 'leek'
c.	s[o:]no '(I) am'	s[o]nno 'sleep'
d.	part[í:] '(I) left'	part[í] '(s/he) left'
e.	v[e:]ro 'true'	v[e]ritá 'truth'
f.	b[ɛ:]ne 'well'	b[ɛ]nde 'bandages'
g.	pres[u:]mo '(I) presume'	pres[u]nto 'presumed'

None of the examples reported in (17), however, constitutes a minimal pair. In cases like b., c., f. and g. the contrast is related to syllable structure, in that the vowel is long when it appears in an open syllable, short when it appears in a closed syllable. In cases like a. and e., the contrast is related to stress placement, in that the vowel is long when it is stressed, short when it is unstressed. Finally, in cases like d. the apparent length contrast is due to the presence of a morphological boundary, i.e. the long vowel is formed by the sequence of two identical vowels brought together by morphological derivation (see contraction in 3.2.2.5; also see Chierchia 1986:6).

According to most scholars (cf. Muljačić 1972:70-2, Serianni 1989:14, Lepschy and Lepschy 1977, Mioni 1993:123), in fact, the opposition between long and short vowels in Italian is not relevant from a phonological point of view, i.e. vowel length is not distinctive and is always predictable from the context. Within the same morpheme, only vowels which are a) in a stressed position, b) in a syllable-final position (i.e. in an open syllable), c) in a non-word-final position are long; all others are short (Mioni 1993:123). In other words, vowels are only long when they feature in a stressed open syllable which is not word-final, or when they span a morpheme boundary.

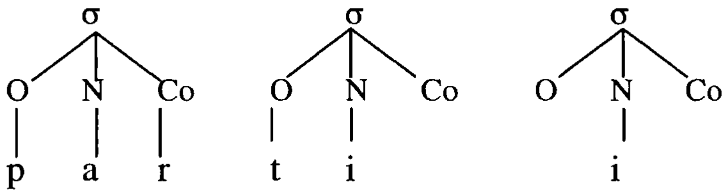
Chierchia (1986:15) actually suggests that bimorphemic long vowels such as the ones in *partíi* '(I) left' and *piíssimo* 'extremely pious' should be represented as a sequence of two vowels (18a), rather than as one vowel linked to two timing slots (18b). The fact that, in a word like *piíssimo*, it is the second segment – and not the first, as is generally the case with long vowels – that is stressed, seems to confirm this hypothesis (ibid.).

(18)

In Italian, according to Chierchia (ibid.), the configuration in (18b) would thus be disallowed in the lexical component.

Although the whole discussion has an undisputable validity, the central point is one that is only understood and never made explicit throughout Chierchia's argumentation: bimorphemic vowels, both at the lexical and at the post-lexical level, are heterosyllabic, i.e. each of their two segments belongs to a different syllable. The diagram in (19) shows that a word like *partii* [par.ti.i] is trisyllabic, because the apparent long vowel is made up of two identical short vowels belonging to separate syllables.

(19)



Notice that, in (19), the onset of the last syllable is left empty. This is a relatively common configuration in Italian, where onsets are not obligatorily filled. In an OT perspective, this means that in Italian the ONSET constraint is ranked rather low. More precisely, it must be dominated by an alignment constraint requiring the right edge of a stem to coincide with the right edge of a syllable, which can be called ALIGN (Stem, R, σ, R). As is shown in tableau (20), ONSET must also be outranked by DEP-C-IO (in order to avoid Consonant Insertion, as in candidate c.) and by MAX-V-IO (in order to avoid Vowel Deletion, as in candidate d.). GF is instead ruled out by the above-mentioned ALIGN (Stem, R, σ, R) (see candidate e.).

(20)

Input: /partI + I/	ALIGN (Stem, R, σ, R)	DEP-C-IO	MAX-V-IO	ONSET
a. par.ti:	*!			
b. ☞ par.ti.i				*
c. par.ti.ji		*!		
d. par.ti			*!	
e. par.tji	*!			

The ranking of ALIGN (Stem, R, σ, R), DEP-C-IO and MAX-V-IO with respect to one another is not relevant here, because each of the sub-optimal candidates violates one of them.

3.2.2 Alternations

Apart from the one in (17d), which has been shown to display a sequence of two identical heterosyllabic vowels rather then a long vowel, the alternations in (17) can all be accounted for on the basis of stress assignment and syllable structure constraints. Although this is a cross-linguistically common feature, it is particularly noticeable and clear-cut in a syllable-timed language like Italian, where syllables all have to be of equal length. As was stated in 3.1, this triggers 'temporal compensation', by which a vowel in an open syllable is longer in order to compensate for the lack of a Co consonant. This applies, however, only to stressed vowels, since a vowel in an unstressed syllable is always short independently of the presence or absence of a Co consonant.

3.2.2.1 Vowel Lengthening

NV (131-2), assuming that all Italian vowels are underlyingly short, postulate a rule of Vowel Lengthening (VL), which 'lengthens the vowel of a strong syllable which is the DTE of a word (i.e. the vowel with primary word stress) iff it is in an open syllable which is not word final'.[15] Furthermore, they argue that the domain of application of the rule is the PW, on the following grounds:

a) The rule applies only to one element in any sequence stem+suffix, which in Italian constitutes a single PW (e.g. *abbaiava* [,abba'**ja:**va] '(it) was barking' < /abbaIa + va/).
b) The rule applies in disyllabic prefixes – which in Italian make up independent PWs - but not in underived words that begin with similar syllables (e.g. *paramilitare* ['**pa:**ra,mili'**ta:**re], in which the prefix *para-* is still recognizable as an independent morpheme, vs. *parametrico* [,para'mɛ:triko], in which it is not).
c) The rule applies twice when the appropriate phonological environment is met in both members of a compound, which in Italian are separate PWs (e.g. *dopolavoro* ['**do:**pola'**vo:**ro] 'after-work activities' < *dopo* 'after' + *lavoro* 'work').

[15] It could be argued, on the contrary, that vowel length is contrastive, and that long vowels are stressed because heavy syllables attract stress. In fact, with the only apparent (see below) exception of word-final stressed vowels, all vowels that bear prominence are either long or in closed syllables, i.e. they are in heavy (bimoraic) syllables. However, since within the same root long vowels can alternate with short vowels in open syllables depending on whether they are stressed (e.g. [ra'dʒ**o:**ne] 'reason' ~ [radʒ**o**'na:re] 'to reason'), one should assume a process of vowel shortening which has no independent basis.

The rule of VL is thus formulated as in (21).

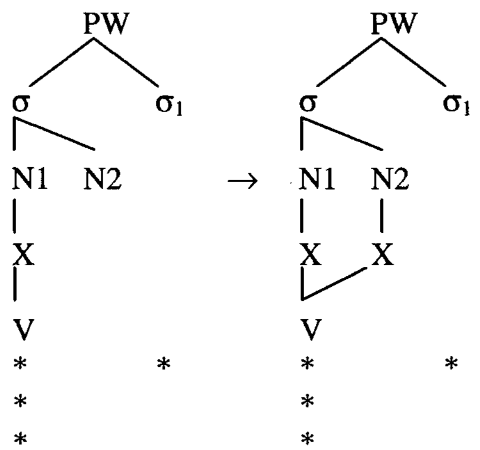
(21) $V \rightarrow [+long] \; / \; [\; \ldots \; [\; \ldots \; \underset{[+DTE]}{\underline{\qquad}} \;]_{\sigma} \; X \;]_{PW}$

(NV:132)

Clearly such a rule, dealing with elements such as segmental length and syllable structure, is better formulated in autosegmental terms: as was mentioned in 3.1, the use of the feature [± long] is highly unsatisfactory. Moreover, the need to specify the presence of stress calls for the application of the metrical grid.

A more accurate - although perhaps not more economical - representation of the above phenomenon is therefore the one in (22).

(22)



where σ_1 = one or more syllables

In Italian, VL in stressed open syllables seems to apply postlexically. If it took place within the lexicon, truncated forms such as those in (23a) would exhibit a long vowel, just like their base forms, reported in (23b).

(23)

(a) *Truncated forms*	(b) *Base forms*
p[a]n pepato 'gingerbread'	p[a:]ne 'bread'
f[a]rlo 'to do it (m.)'	f[a:]re 'to do'
ven[i]r via 'to leave'	ven[i:]re 'to come'

s[o]l leone 'summer heat' s[o:]le 'sun'

M[a]r Morto 'Dead Sea' m[a:]re 'sea'

Truncation must in fact be a post-lexical phenomenon, in that it is triggered by the presence of a following word or clitic. If VL applied in the lexicon, it would precede truncation and, since the non-truncated forms have stressed vowels in open syllables, it would lengthen such vowels. It is only after truncation that the consonants following the stressed vowels are resyllabified into the preceding syllable, thus closing it. VL cannot apply precisely because the open syllable environment is not met, and must therefore be ordered after truncation, i.e. post-lexically. According to Chierchia (1986:23), this is a consequence of the fact that long vowels are disallowed in the lexical component (see 3.2.1 above).

So far we have assumed, with NV, Chierchia (1986) and many others, that long vowels are absent from the URs of Italian words, and that they only surface in the post-lexical component. This assumption is in keeping with the widely accepted notion that long vowels are cross-linguistically marked. Some scholars, however, seem to look at Italian vowel length alternations from the opposite point of view, i.e. assuming a shortening process. Serianni (1989:14), for example, mentions that Italian vowels are short in the following three cases: a) when they are unstressed, b) when they are stressed but appear in a closed syllable, c) when they are stressed and appear in an open syllable but are word-final (cf. *salí* [sa'li] '(s/he) went up', *salto* ['salto] 'jump' vs. *sale* ['sa:le] 'salt'). However, while the definition of an environment for Vowel Shortening involves a three-way disjunction, the environment of VL is much more easily circumscribable: a vowel is only long when it is in a non-final stressed open syllable. On the basis of this observation, and of cross-linguistic generalizations, we will continue to assume the absence of long vowels from the Italian lexicon.

Within an OT framework, VL in stressed open syllables is triggered by the constraint SWP, requiring stressed syllables to be heavy and thus banning monomoraic stressed syllables. This must outrank DEP-μ-IO (or DEP-X-IO), which militates against lengthening (cf. 3.1). Clearly, *V: must be ranked low, too.

The tableaux for the pair *bene* ['bɛ:.ne] 'well' / bende ['bɛn.de] 'bandages' are reported in (24) and (25) respectively.[16]

(24)

Input: /bɛne/	SWP	DEP-μ-IO	*V:
a. 'bɛ.ne	*!		
b. ☞ 'bɛ:.ne		*	*

In (24), the optimal candidate is b., in that it satisfies SWP, though violating the lower ranked DEP-μ-IO and *V:.

(25)

Input: /bɛnde/	SWP	DEP-μ-IO	*V:
a. ☞ 'bɛn.de			
b. 'bɛ:n.de		*!	*

In this second case, the selected candidate is the one with the short vowel, since the stressed syllable, being closed, is already heavy, thus satisfying the SWP without any need for violating DEP-μ-IO (or *V:).[17]

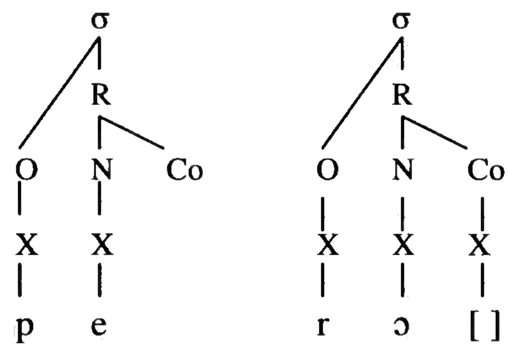
3.2.2.1.1 Lack of VL in word-final position

In Italian, as has been mentioned several times, stressed vowels that are in absolute word-final position are never lengthened (cf. *cosí* [ko'zi] 'so', *peró* [pe'rɔ] 'but', *vanitá* [vani'ta] 'vanity'). This curious phenomenon has been explained in different ways. Perhaps the most popular argument (see, among others, Saltarelli 1983, Selkirk 1984a, Chierchia 1986, NV) is that Italian word-final stressed vowels do not, as it seems, appear in an open syllable, but in a closed syllable with an unfilled position in the coda; in other words, that all word-final stressed syllables are closed, so that the environment for VL is never met in such a position. A word like *peró* will therefore be represented as in (26).

[16] For the time being, the problematic lack of lengthening in word-final stressed open syllables has been left aside (but see 3.2.2.1.1).

[17] Notice that candidate b., besides violating DEP-μ-IO and *V:, also violates *3μ, since it exhibits a trimoraic syllable.

(26)



This representation has its roots in the history of Italian: in this language, oxytones are generally derived from consonant-final Latin words, through the deletion of the last consonant (e.g. Lat. *per hoc* > It. *peró*, Lat. *vanitas* > It. *vanitá*). At the synchronic level, as was mentioned in 1.4, this argument is based especially on the facts related to the process of Syntactic Doubling, whose reduplicated consonant is construed as filling in the empty coda position of the preceding syllable (cf. Saltarelli 1983:16, NV:165-168). With regard to this, Chierchia (1986:9) points out that

> While word internally vowel length and consonant length are both possible ways of meeting well-formedness (i.e. of rendering heavy stressed syllables), only consonant length appears to be possible at word boundaries.

He (ibid.:25) explains this fact by extending the lexical ban on long vowels (see above) to certain 'close knit' phrasal environments. This would be a consequence of resyllabification within 'strict intonational units' behaving in virtually the same way as word-internal syllabification. Without resorting to the notion of 'close knit phrasal environment', however, it is perhaps sufficient to postulate that, since the final slot in words like *peró* [pe'rɔ[ ]] or *vanitá* [vani'ta[ ]] is left empty by the diachronic deletion of a consonant, it must be a C-slot, which can only be filled with a [+consonantal] segment.

The OT analysis of a word like *peró*, if this is assumed to have a final empty slot in the input, is unproblematic, in that it is analogous to that of *bende* (see tableau (27)).

(27)

Input: /perɔ[ ]/	SWP	DEP-μ-IO	*V:
a. ☞ pe'rɔ[ ]			
b. pe'rɔ:[ ]		*!	*

An alternative solution, which has the advantage of avoiding the postulation of an empty slot in the input, is the use of the independently justified constraint *FINAL-VOC: (see 2.3), disallowing long vocoids in word-final position. If this constraint is ranked above SWP, word-final vowels will be short even when stressed. The alternative tableau of *peró* is given in (28).

(28)

Input: /perɔ/	*FINAL-VOC:	SWP	DEP-μ-IO
a. ☞ pe'rɔ		*	
b. pe'rɔ:	*!		*

Still within an OT perspective, Bullock (1998:61-3) accounts for this phenomenon in a rather different way. She assumes that SWP[18] is outranked by the alignment constraint Align(LxWd, R, Segment, R), requiring that a lexical word be aligned with a segment at its right edge. This ranking leads to a preference for a monomoraic stressed syllable (violating SWP) over a bimoraic syllable with an empty mora[19] (which would be phonetically interpreted as a lengthened vowel) at the right edge of the lexical word (cf. the tableau of *giá* ['dʒa] 'already' in (29)).

(29)

Input: /dʒa/	Align R(LxWd, segment)	SWP
a. ☞ $[(dʒa)]_{LxWd}$		*
b. $[(dʒ[a]_{\mu}[\ ]_{\mu})_F]_{LxWd}$	*!	

As was mentioned before, while lengthening does not take place at a word boundary, Syntactic Doubling does. This can be explained, according to Bullock (ibid.:62), on the basis of the latter phenomenon referring to PW, rather than Lexical Word (LxWd)[20], boundaries. In Italian, these two constituents do not coincide: while the LxWd is aligned with a segment, the PW is aligned with a foot. Therefore, a word like *già* will be parsed as $[(dʒ[a]\ |_{LxWd}\ [\ ]_{\mu})_F]_{PW}$. The non coincidence of LxWd and PW clearly implies a violation of the constraint LxWd = PW (GW = PW in 1.5).

[18] Bullock (1998) actually uses the constraint WSP, by which heavy syllables must be stressed, rather than the constraint SWP, by which stressed syllables must be heavy. In this context, however, reference to the latter seems more appropriate, so that 'WSP' has been replaced with 'SWP' throughout the account. WSP becomes instead relevant when explaining the preference for stress assignment to CVC syllables.
[19] In Bullock's (1998) analysis, the mora is not only a unit of weight, but also a prosodic constituent.
[20] Bullock's (1998) LxWd is what was called a 'Grammatical Word' in 1.5.

3.2.2.1.2 Consonantal and vocalic length

Strictly dependent on syllable structure is also the issue of the relationship between consonantal and vocalic length in Italian. Contrary to vowel quantity, consonant length is generally considered to be distinctive in this language. Minimal pairs such as *pala* 'shovel' ~ *palla* 'ball', *rene* 'kidney' ~ *renne* 'rein-deer (pl.)' or *moto* 'movement' ~ *motto* 'aphorism', in which the difference in spelling corresponds to an actual difference in consonant-length, are extremely common in the Italian lexicon. At the phonetic level, however, such pairs are distinguished not only by consonant length, but also by vowel length (cf. (17b) and (17c) above): whenever a consonant is long, the preceding vowel is short (as in *palla* [palla]), and vice-versa (as in *pala* [pa:la]), unless the consonant is followed by another consonant (as in *palma* [palma] 'palm-tree'). This phenomenon only applies to stressed vowels, and corresponds to what Saltarelli (1983:9-10) calls 'Duration Rhythm'. It is clearly a corollary to the VL facts described above: just like the [l] sound in *pal.ma*, the first segment of the geminate in words like *pal.la* or *ren.ne* belongs to the rhyme of the syllable in which the relevant vowel is placed, thus closing the selfsame syllable. As a consequence, the vowel must be short, since vowels are always short in closed syllables in Italian.

By looking at these examples, one could actually argue that it is not vowel length that is non-contrastive and entirely predictable from the context, but consonant length: whenever a vowel is long, the following consonant must be short, whenever it is short, the following consonant must be long. However, this would anyway require the further condition that long vowels cannot surface in closed syllables (i.e. an independent rule of CSS), in order to account for forms like *palma* [palma]. Therefore, the analysis positing vowel length as phonetic and consonant length as phonemic appears to be more economical.

Notice, moreover, that, while long vowels only appear in stressed syllables, long consonants do not only follow stressed vowels, but also unstressed vowels (e.g. *cappone* [kap.'po:.ne] 'capon', *terreno* [ter.'re:.no] 'ground', *collante* [kol.'lan.te] 'glue', *ferroso* [fer.ro:.zo] 'ferrous'. This is a consequence of the fact that, while stressed syllables are obligatorily heavy, unstressed syllables may be either heavy or light.

In the Northern varieties of Italian, however, double consonants often undergo degemination. There is actually a tendency, due to the influence of the dialectal 'substratum', for degemination to become fossilized, so that, for certain speakers at least, Northern Italian does not exhibit geminate consonants at all. In such cases, two words like *fato* [fa:to] 'fate' and *fatto* [fato] 'fact; made' are distinguished merely on the basis of vowel length, so that such a feature becomes contrastive. The analogy with languages such as French or English naturally springs to mind. Diachronic degemination, however, is by no means completed, and it is doubtful that it will ever be, due to the influence of the written standard. Application of this process is still heavily dependent on extralinguistic variables such as social class, register or emphasis. For most speakers, and especially for the vast majority of the younger ones, who do not display active knowledge of the dialects, double and single consonants are simply in free variation, so that both [fatto] and [fato] are accepted as pronunciations of <fatto>. Moreover, the latter pronunciation often happens to be stigmatized, at least in certain contexts, so that it becomes a mere diaphasic variant. The resistance of geminates and the consequent failure of vowel length to become contrastive appear to be due mainly to the fact that, while consonant duration is graphically represented, vowel duration is not. On the whole, degemination in Northern Italian can be regarded as a synchronic phenomenon, so that vowel length cannot really be seen as distinctive. In OT terms, since geminate consonants are cross-linguistically marked, there seems to be a constraint disallowing them, which we may call *GEM. In Northern Italian, such a constraint must outrank SWP, since stressed open syllables are light (see tableau (30)). Clearly, for those speakers who exhibit free variation between double and simple consonants, free ranking of the two constraints must be postulated.

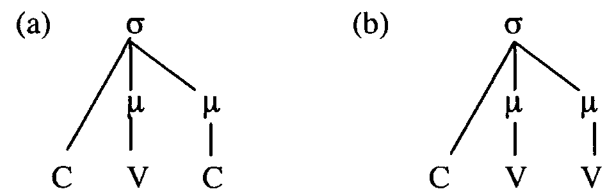
(30)

Input: /fato/	*GEM	SWP
a. ☞ fa.to		*
b. fat.to	*!	

3.2.2.2 A moraic analysis

As was mentioned in 1.3 and 3.1, the relationship between vowel length and syllable structure can be effectively represented in terms of moras, as it is evidently to do with syllable weight. Since both a Co consonant and the second segment of a long vowel or diphthong (i.e. an N2-vocoid) are assigned one mora, both (C)VC and (C)VV syllables will be bimoraic, as is shown in (31).[21]

(31)



Since we are dealing with length within the rhyme, referring to moras here becomes essential, in that onset segments, which are not assigned a mora but still have a skeletal slot, are not taken into account. This explains the difference between a word like *madre* [ma:.dre] 'mother', in which the stressed vowel is lengthened, and a word like *manda* [man.da] 's/he sends', in which the stressed vowel remains short. Although the number of intervocalic C-slots is the same in both words, the first consonant of the cluster in *manda*, being syllabified in the coda, is assigned a mora, whilst in *madre* no consonant is moraic because they are all in the onset.

Italian, like many other languages, has maximally bimoraic syllables (cf. (9), repeated here as (32)).

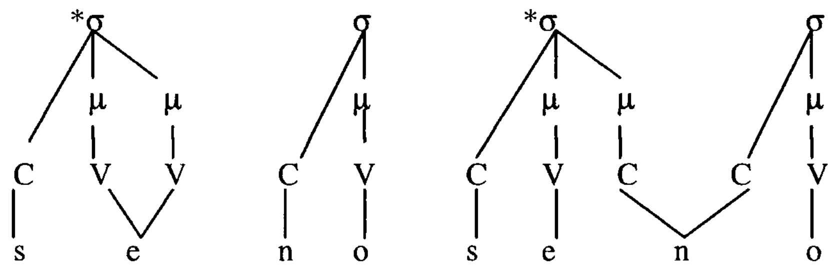
(32) $\sigma \rightarrow \mu(\mu)$

At least in native words, moreover, it does not appear to have extrasyllabic consonants, so that syllables can only be of the types (C)V (monomoraic), (C)VV and (C)VC (both bimoraic). Stressed syllables, on the other hand, can only be bimoraic, i.e. either they are closed by a coda consonant, or they display a long vowel.

Duration Rhythm, therefore, is easily accounted for in terms of moras: the double contrast between two lexemes like *seno* [se:no] 'breast' and *senno* [senno] 'wisdom' can be formulated as in (33).[22]

[21] For the sake of brevity, syllabic structure has been omitted.

(33)



It should be noted that the number and structure of the syllables in each word is exactly the same, but that, while in the first word the second mora is attached to the second segment of a long vowel, in the second word it is attached to the first segment of the geminate consonant. Nevertheless, it must be borne in mind that, in this framework, the mora is not regarded as a constituent of the syllable, but as a mere unit of weight. In fact, if the internal structure of the syllables is taken into consideration, the second segment of the long vowel in *s[e:]no* will be placed in the nucleus of the first syllable, while the first segment of the geminate consonant in *se[nn]o* will be placed in its coda.

3.2.2.2.1 Bimoraic Enforcement and the uneven trochee

In Italian, syllables are not only maximally, but also optimally bimoraic. This is called by Prieto i Vives (2000) *Bimoraic Enforcement*, and it applies with different strength in stressed and unstressed syllables: while the latter may be monomoraic, stressed syllables are obligatorily bimoraic. In fact, Chierchia (1986:8; 22) refers to this restriction as to the *Strong Rime Condition*, by which 'stressed syllables (in non prepausal position) must be heavy'. Notice how this language-particular condition resembles very closely the above-mentioned universal constraint SWP: while in DT a restriction has to be specified as present or absent in each language, in OT the language-particular flavour is given by the higher or lower ranking of the constraint. In Italian, such a constraint must be ranked high. It must dominate not only, as was said above, DEP-μ-IO, but also the constraint demanding evenness in trochaic feet, namely EVEN-TROCHEE. As was mentioned in 3.1, this latter reflects the cross-linguistic preference for the even trochees (LL) and (H) over the uneven trochee

[22] Where '*' indicates PW-stress.

(HL). Mester (1994, in Prieto i Vives 2000) draws up a scale of foot well-formedness in trochaic languages, which is reported in (34).

(34) *Trochaic foot well-formedness scale:*

bimoraic trochees ($\sigma_{\mu\mu}$ or $\sigma\sigma$) ≻ trimoraic trochees ($\sigma_{\mu\mu}$ σ) ≻ monomoraic trochees (σ)

In a trochaic language like Italian, however, the dominant foot type is the uneven (HL). Bullock (1998:59) points out that, just like in French, the bimoraic minimum is predicated of the syllable rather than of the foot. The large presence of (HL) feet in Italian can be accounted for precisely by ranking SWP above EVEN-TROCHEE. As the tableau of *mare* 'sea' in (35) shows, the requirement for stressed syllables to be heavy results in HL sequences; its action, combined with that of another high-ranked constraint disallowing unparsed syllables (PARSE-SYL), produces uneven trochaic feet, which violate the lower-ranked EVEN-TROCHEE.

(35)

Input: /mare/	SWP	PARSE-SYL	EVEN-TROCHEE
a. ['mare]$_F$	*!		
b. ['ma:]$_F$ re		*!	
c. ☞ ['ma:re]$_F$			*

Bullock (ibid.:65-8), however, argues that, although EVEN-TROCHEE is ranked low in Italian, its presence as a universal constraint is still felt in the tendency to eliminate uneven trochees on some occasions. Rather than doing this by making the heavy syllable in (HL) light, thus obtaining an (LL) foot, Italian eliminates the second (light) syllable altogether, so that a monosyllabic – but evenly bimoraic, and thus well-formed - (H) foot arises.[23] This is what happens with truncated and reduplicated forms, which are both accomplished by eliminating the post-tonic vowel, and thus the second syllable of the uneven trochee, whenever possible - i.e. whenever the modified coda filter demanding that a coda consonant be a sonorant (see 1.3) is fulfilled. Some examples are given in (36).

[23] According to Bullock (1998), the absence of (LL) feet in languages like Italian and French is due to the fact that 'foot based conditions on prosodic size and wellformedness [...] have been yielded over to syllable based conditions'.

(36)

(a) *Truncations* (see also (23) above):

cammino > cammin 'path, walk'

grande > gran 'big'

andare > andar 'to go'

vuole > vuol '(s/he) wants'

(b) *Reduplications*:

belbello (< bello 'pretty') 'slowly'

pianpiano (< piano 'slowly') 'gradually'

orora (< ora 'now') 'just now'

benbene (< bene 'well') 'thoroughly'

(from Bullock 1998:66)

The sonorant consonant is preserved, rather than being deleted together with the following vowel, because of the constraints MAX–BT and MAX-BR, requiring the truncated form and the reduplicant respectively to correspond in every element to the base. The MAX constraints, however, will have to be outranked by EVEN-TROCHEE, in order for the post-tonic vowel to be absent from the truncated/ reduplicated forms.
One last observation is that, since CVC syllables are bimoraic, the constraint requiring coda consonants to be moraic (namely WEIGHT-BY-POSITION) must be ranked high in Italian. *FINAL-C-μ, by which word-final consonants are not moraic, must instead be ranked very low.

3.2.2.3 Length in trisyllabic feet

Although Italian is said to be a syllable-timed language, the length of its stressed vowels is also related to foot – or, in Donegan's (1985) words, 'stress-group' - structure. In fact, a stressed vowel in an antepenultimate open syllable is noticeably shorter than a stressed vowel in a penultimate open syllable. This phenomenon is exemplified in alternations such as *cono*['kɔ:no] 'cone' ~ *conico* ['kɔˑniko][24] 'cone-shaped', *vedo* ['ve:do] '(I) see' ~ *vedono* ['veˑdono] '(they) see'.

The same type of alternation can be found in words in which the stressed vowel, being in a closed syllable, is considered to be short, as in *vendo* ['veˑndo] '(I) sell' as opposed to *vendono* ['vendono] '(they) sell'. The general tendency for the stressed syllable in trisyllabic feet to be shorter than that in disyllabic feet is the same tendency that underlies Trisyllabic Shortening (see 3.1). It is a consequence of the phenomenon of stress-group (or foot) isochrony, by which the temporal distance between two stresses remains constant even if the number of prosodic units (i.e. syllables) comprised in that space varies. As was mentioned in 3.1, such a phenomenon has been said to affect mainly stress-timed languages. From the facts described above, however, it follows directly that it can and does apply also in syllable-timed languages like Italian.[25]

It has already been said (3.1) that in many stress-timed languages syllable isochrony takes place alongside stress-group (foot) isochrony. By the same token, it is apparent that syllable-timed languages can exhibit stress-group (foot) isochrony alongside syllable isochrony. This seems to indicate that the distinction between the two types of language is somewhat blurred, and that Nespor (1993:258-9) has a point in rejecting such a distinction altogether (cf. 1.4).

An OT analysis of this phenomenon involves the combined action of the constraints FT-BIN – demanding that feet contain either two moras or two syllables – and PARSE-SYL, by which all syllables must be parsed into feet. As tableau (37) shows, the correct output for a form like [veˑdono] is obtained by ranking these two constraints above SWP.

[24] The half-lengthening of the stressed syllable is probably a mere phonetic effect.

[25] Sluyters (1990:81, in Bullock 1998:59) argues that absence of vowel lengthening in ante-penultimate stressed open syllables is found only in certain varieties of Italian. According to Bullock (ibid.), who does not accept ternary feet, this would indicate that such varieties have (LL) even trochees alongside (HL) uneven trochees. The ternary foot interpretation, however, seems to be less stipulatory and more phonetically grounded.

(37)

Input: /vedono/	PARSE-SYL	FT-BIN	SWP	DEP-μ-IO
a. [ve:dono]$_F$		**!		*
b. ☞ [vedono]$_F$		*	*	
c. [ve:do]$_F$<no>	*!			*
d. [vedo]$_F$<no>	*!		*	

The double violation of FT-BIN is fatal to quadri-moraic [ve:dono] with respect to trimoraic [vedono] (phonetically [veˑdono]). In order to rule out candidates c. and d., moreover, PARSE-SYL must dominate FT-BIN.

3.2.2.4 Other lengthening processes

3.2.2.4.1 Final Lengthening

As in most other languages, Italian, too, exhibits prosodic Final Lengthening (FL). As was said in 3.1, this is a phenomenon by which the DTE of the last PW within a given phrasal constituent is lengthened. The nature of this constituent varies from one language to another: NV (176) establish that in Italian, just as in English, the PPh is the domain of application for this phenomenon.

More precisely, Italian FL is a PPh limit rule of rhythmic readjustment. The lengthening at the end of a PPh, in the case of a word-final stressed syllable, introduces enough distance from an immediately following peak as to block Stress Retraction (SR, see 1.4), which is in fact a PPh span rule (NV:174-7). In the example in (38), for instance, the stress clash between the last syllable in *veritá* 'truth' and first in *sálta* 'jumps, comes' is not resolved by means of SR, because the PPh-final /a/ of *veritá* is lengthened.

(38) [la veritá]$_{PPh}$ [sálta fuori]$_{PPh}$ [quasi sempre]$_{PPh}$

'The truth almost always comes out'

(ibid.:175)

Nespor (1993:68) shows how FL can also act as a clue to syntactic disambiguation when the same string of words may in principle be interpreted in two different ways. The sentence in (39a), for example, is to be interpreted either as in (39b) or in (39c), according to whether the foot [cása]$_F$ or [sólo]$_F$ is lengthened, respectively.

(39)

(a) Luca stava in casa solo se pioveva.

(b) Luca stava in c[a:]sa]$_{PPh}$ solo se pioveva.

'Luca stayed at home only if it rained'

(c) Luca stava in casa s[o:]lo]$_{PPh}$ se pioveva.

'Luca stayed at home alone, if it rained'

IPh-final lengthening can also be found in Italian. Again, it is the last foot that is lengthened. As can be seen in (40), the same foot [vesse]$_F$ is longer in (40a) than in (40b), although it is placed at the end of a PPh in both cases.

(40)

(a) Federica, [nonostante pio**v[eˑ]sse**]$_{IPh}$, andó a casa a piedi.

'Federica, although it rained, went home on foot'

(b) [Se pio**v[e]sse** mi fermerei]$_{IPh}$.

'If it rained I would stop'

3.2.2.4.2 Compensatory Lengthening

There seems to be no Compensatory Lengthening (CL) in Italian, and this goes to confirm De Chene's (1985) hypothesis that CL can only take place in languages exhibiting contrastive vowel quantity (cf. 3.1).

3.2.2.5 Contraction

As was mentioned in 3.2.1, apparent long vowels in words such as *salii* 'I went up' and *piissimo* 'extremely pious' actually exhibit a sequence of two heterosyllabic identical vowels, because the two segments belong to separate morphemes. However, not all heteromorphemic sequences of identical vowels behave in this way. The [II] clusters in the words in (41) undergo a process of contraction, by which the two segments are reduced to one.

(41)	vari + i	>	varii	[va:ri:]	'varied (pl.)'
	armadi + i	>	armadii	[arma:di:]	'wardrobes'

The difference between the two types of lexical items corresponds to a difference in stress assignment. In words such as *varii* or *armadii*, which undergo contraction, neither of the two identical vowels is stressed at any point in the derivation (cf. the singulars *vário* 'varied (sg.)' and *armádio* 'wardrobe'). In words like *salíi* '(I) went up', on the other hand, the first of the two identical vowels is stressed (cf. the infinitive *salíre* 'to go up'), thus blocking contraction. *Piíssimo* 'extremely pious' is refractory to this process for exactly the same reason; although its root-final vowel does not surface as prominent, we can assume that it was so at some point in the derivation, since *pío* 'pious' is stressed on its first vocoid. It is not the stress on the second (i.e. the suffixal) vocoid that blocks contraction, because the process does take place in *varíssimo* 'extremely varied'.

A brief OT analysis of these facts is as follows. In 2.2, it was pointed out that the constraint PK-MAX-IO – by which if a segment is the stress peak of the input its correspondent must be the stress peak of the output – is ranked high in Italian. This is what blocks contraction of a stressed vowel with another vowel, and must therefore be ranked above the contraction triggering constraint NOHIATUS (see 3.1). The tableau for *salíi* is reported in (42).

(42)

Input: /sa'lI + I/	PK-MAX-IO	NOHIATUS
a. ☞ sa.li.i		*
b. sa.li	*!	
c. sa.lji	*!	

Notice that, as was said in 2.2, this ranking also blocks GF.

3.3 Long and short vowels in Arabic

This section will deal with vowel quantity in Arabic at large and in AA in particular. I will first discuss the quality of AA long vocoids and the way they should be represented, then I will tackle the issues of phonemic length and of the factors conditioning vowel length. The second part of this section will be about alternations in vowel length; it will first deal with shortenings and then with lengthenings, touching briefly on Vowel Contraction in the last subsection. As always, an alternative OT account of the phenomena will also be given.

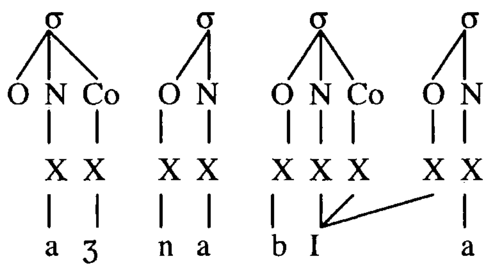
3.3.1 Representation of vocoid length in AA

Phonetically, vowel (and consonant) length contrasts are a particularly marked feature in most Arabic dialects, where quantitative rhythm plays a fundamental role (Cantineau 1960:94-5): the length of long vowels is often more than double the length of their short counterparts (see for example Al-Ani 1970:75). In fact, traditionally – and in the moraic model – a short vowel is assigned one mora, a long vowel two moras (Hayes 1989, 1995).

All three phonological vowels (i.e. /a/, /U/, /I/) can be long or short, both in Ammani Arabic (AA) and in Standard Arabic (SA). In addition to [a:], [u:] and [i:], however, AA displays the long vowels [e:] and [o:] which, as was mentioned in 2.3, are derived from /aI/ and /aU/ respectively, by means of monophthongization.[26] By the same token, Cantineau (1960:102) points out that, in Classical Arabic, even [i:] and [u:] should not always be considered as long [i] and [u] respectively, since they often correspond to *–iy-*, *-uw-* or *–yi-*, *-wu-* (e.g. *kiis* 'sack' < *kiys* (pl. *'akyaas*), *suuq* 'market' < *suwq* (pl. *'aswaaq*), *yamuutu* 'he will die' < *yamwutu*, *tabiiDu* 'she will lay eggs' < *tabyiDu*). This analysis can be applied to the modern Arabic dialects, where similar alternations take place. In a framework which does not differentiate between glides and the homorganic vowels underlyingly, these forms should still be distinguished from phonemically long vowels, since one vocoid (the glide) belongs morphologically to the root tier, while the other vocoid (the vowel) belongs to the affix tier. Despite the fact that, as a consequence of tier conflation and of syllable structure assignment, their surface realization is identical to that of a long vowel, and is thus to be transcribed as a single vowel linked to two slots (in short, [V:]), their UR is a sequence of two separate vocoids (in short, /VV/).

Although the following discussion will focus on the length of nuclear elements, it is important to remember that, in the Arabic dialects, glides can also be doubled. Clearly, in our framework double glides are underlyingly undistinguished from long vowels. Just like geminated contoids, a high vocoid can span the coda of one syllable and the onset of the next (e.g. ***aw.wal*** 'first'), or even cover three successive positions as in the case of the feminine of 'nisba' adjectives (see 2.3), where the final vocoid of the masculine is lengthened by one timing slot to fill in the empty onset of the suffixal syllable /-a/. It is clearly a case of Glide Insertion (see 5.4).

(43) /aʒnabI:/ + /a:/ → [ʔaʒnab**ij.j**a] 'foreign, foreigner (f.)'



3.3.1.1 Phonemic length

In AA, as in SA (see Al-Ani 1970:22) and most other dialects (see Harrell (1957:52) for Egyptian, Abumdas (1985:45) for Libyan Arabic), vowel length is contrastive. As far as AA is concerned, this is clearly illustrated by the minimal pairs in (44), where each of the long vowels is contrasted with the homorganic short vowel and with each of the other long vowels.

(44)	*f[a]laH* 'to cultivate'	~	*f[a:]laH* 'successful'
	9[a]lam 'flag; figure'	~	*9[a:]lam* 'world'
	s[i]nn 'tooth'	~	*s[i:]n* 's'
	j[a]rr 'to pull'	~	*j[a:]r* 'neighbour'
	k[u]ll 'all, every, each'	~	*k[u:]l* 'eat! (m.sg.)'[27]
	j[a:]b 'he brought'	~	*j[i:]b* 'bring! (m.sg)'
	r[a:]H 'he went'	~	*r[u:]H* 'go! (m.sg.)'
	d[i:]r 'monastery'	~	*d[u:]r* 'turn! (m.sg.)'

Therefore, vowels (and segments at large) are underlyingly long or short, and cannot be underspecified for length in phonological representations. Vowel quantity is only

[26] As was mentioned in 1.1, North African dialects display a wider inventory of vocoid phonemes, including long vowels, but this is not of concern here.

partially conditioned by prosodic factors such as syllable structure and stress assignment. In AA, as in many other Arabic varieties, both short and long vowels may appear both in stressed and unstressed, in open and closed syllables, as can be seen from the examples in (45).

(45) *Short vowels:*

Stressed in closed syllables: *áH.mar* 'red', *kíl.ma* 'word', *ghúr.fa* 'room'
Stressed in open syllables: *xá.shab* 'wood', *mí.thil* 'like', *rú.bu9* 'quarter'
Unstressed in closed syllables: *ban.dúura* 'tomato', *is.tán.na* 'wait!', *bur.tu.qáan* 'oranges'
Unstressed in open syllables: *Ta.bíib* 'doctor', *si.gáa.ra* 'cigarette', *hu.náak* 'there'

Long vowels:

Stressed in closed syllables: *jáam.9a* 'university', *fáah.ma* 'understanding (f.sg.)'
Stressed in open syllables: *Ha.wáa.li* 'approximately', *síi.di* 'sir', *fúu.Ta* 'towel'
Unstressed in closed syllables: *ta.maan.tá9sh* 'eighteen'
Unstressed in open syllables: *ta.maa.níin* 'eighty', *nii.sáan* 'April', *ruu.máa.ni* 'Roman'

3.3.1.2 Factors conditioning vowel length

3.3.1.2.1 Syllable structure

Despite the data in the previous subsection, long vowels are rarely found in closed syllables word-internally, and the few available examples of such superheavy syllables (McCarthy 1979a, b) are generally only attested in derived environments,[28]

[27] The last three pairs can be considered as minimally contrastive because, in a variety like AA, the geminate consonants in *sinn, jarr* and *kull* have been simplified.

[28] The fact that, in many eastern Arabic dialects, CV:C syllables can be found word-internally in derived environments, while CVCC syllables are only found in word-final position, seems to indicate that the two types of super-heavy syllables are not prosodically identical (Watson 1999b:1). Moreover,

or as a consequence of the operation of phonological processes such as Vowel Deletion (see 4.3). Thus [fa:h.ma] 'understanding (f.sg.)', which preserves the length of its first vowel, is derived from /fa:.hi.ma/ (cf. [fa:.him] 'understanding (m.sg.)'). Rarest of all (and always moot) are examples of long vowels in unstressed closed syllables. According to Holes (1995:56), contrary to the SA of religious readings (Classical Arabic), in Modern Standard Arabic (MSA) long vowels generally appear only in stressed or open syllables (i.e. they are always stressed when in a closed syllable). In the more innovative dialects of Lower Egypt, as will be shown below, CV:C syllables are totally banned from non-final positions, whether they are stressed or not.

A plausible explanation for the above-mentioned facts relates to the Bimoraicity Constraint (cf. 3.1), by which all syllables are maximally and optimally bimoraic, and hence trimoraic syllables such as CV:C are strongly dispreferred, if not altogether banned (Broselow 1992:10). This is a universal constraint on syllable well-formedness, but the strength with which it is implemented varies from variety to variety. In Bahraini (Sunni dialect), for instance, CV:C syllables appear in all positions, as in *jaalbuut* (type of boat) or *yisoolfuun* 'they chat' (Holes 1995:63-4).

Even in dialects which ban domain-initial and –medial superheavy syllables, domain-final CV:C sequences are very commonly found (cf. AA *majnuun* 'crazy', *xanziir* 'pig', *kitaab* 'book'). This can be explained by regarding the final consonant as extrasyllabic, so that the last syllable in each of these words is construed as (regularly) bimoraic (cf. 1.3; 3.1). As was mentioned in 1.4, however, such a syllable itself may be construed as extraprosodic unless followed by the extrasyllabic consonant, in order to account for those dialects in which not only word-final CVC syllables, but also word-final CVV syllables are ignored by stress assignment.[29]

As was said in 3.1, the Bimoraicity Constraint is easily translatable in OT terms, by means of the two constraints $*3\mu$ and $*\sigma_{\mu}$. In particular, $*3\mu$ militates against the appearance of long vowels in closed syllables. The faithfulness constraint MAX-μ-IO, on the other hand, disallows vowel shortening (see below), while DEP-μ-IO disallows

---

'while in most Arabic dialects word-final CV:C and CVCC syllables invariably attract word stress, in San9ani word-final CVCC attracts stress but CV:C attracts stress iff the penultimate or antipenultimate syllable is neither CVV or CVG (i.e.CVGeminate)' (Ibid.).

[29] Since AA does not exhibit word-final long vowels (see below), we will not pursue this issue any further.

vowel lengthening. Extrametricality is instead regulated by the competing constraints WEIGHT-BY-POSITION, requiring the assignment of a mora to all coda consonants, and *FINAL-C-μ, by which word-final consonants should be weightless: in order to obtain extrametrical consonants, the latter constraint must dominate WEIGHT-BY-POSITION. The WSP - requiring heavy syllables to be stressed - bans unstressed heavy syllables, and is doubly violated by unstressed superheavies.

3.3.1.2.2 Foot and word structure

Length is not only related to syllable structure, but also to foot structure - i.e. to stress group isochrony - and to word structure. In Cairene Arabic, for instance, each PW cannot contain more than one long vowel (Holes 1995:50). This claim is supported by the tendency to avoid sequences of long vowels exhibited by dialects in which the ban is not as absolute as in Cairene. For instance, Benhallam (1980:55-6) points out that, in SA, a long vowel may not be created when the preceding syllable contains another long vowel. The example he gives is that in (46), where, after Glide Deletion has applied (see 2.3 and 4.3), the two /i/s do not simply undergo contraction yielding [i:], but are simplified to [i].

(46) /ra:mij+in/ (or /ra:mii+in/) → [ra:min] 'those who throw (m.)'

(ibid.:55)

Although his definition of 'short' and 'long' syllables is not very clear, Benhallam (ibid.:57) has the intuition that the sequences long-long-long (apparently a string of three consecutive heavy syllables) and – especially - short-short-short (apparently a string of three consecutive light syllables) are avoided when possible (cf. 4.3).

Finally, Cantineau (1960:96) observes that the closer a vowel is to the end of the word, the longer it is. This is likely to have something to do with the distribution of word stress and the prominence of the last foot, since stress assignment in Arabic is heavily dependent on syllable weight. As was explained in section 1.4, in SA and the Levantine dialects prominence is straightforwardly assigned to the last syllable if it is superheavy, otherwise to the penultimate if it is heavy, otherwise to the antepenultimate. Since a syllable containing a long vowel is by definition at least heavy, the last long vowel in a word will receive main stress, unless it is in a word-final open syllable, which is just heavy and not superheavy, as required by the stress assignment algorithm. In varieties like AA, even this latter possibility is ruled out, because long vowels never appear word-finally (cf. 3.3.2.2); therefore, the last long vowel is always stressed. Given that accented vowels are phonetically longer than unaccented ones, this explains why the closer a long vowel is to the end of the word, the longer it seems to be.

3.3.2 Quantitative alternations

Although vowel length can be distinctive in Arabic, long and short vowels often alternate. Depending on whether the underlying phoneme is long or short, the alternations can be classified into shortenings and lengthenings, respectively.

3.3.2.1 Shortenings

3.3.2.1.1 Closed Syllable Shortening

Benhallam (1980:41; 52-3), in relation to SA, mentions a process of Vowel Shortening (VS) that shortens long vowels when they are followed by two consonants, yielding alternations such as [jafi:qu] 'he wakes up' ~ [tafiqna] 'you (fem. pl.) wake up'. He formulates the rule of VS as follows:

(47) V: → V / ____ CC

(ibid.:52)

A similar process applies in virtually all Arabic dialects. Benhallam's formulation, however, is flawed, in that it does not take into account syllable structure, which is clearly the factor determining the application of this process. Since, in most Arabic varieties, two consecutive consonants are syllabified into different syllables (see 1.3), the shortening applies consistently to a long vowel in a closed syllable. It is therefore an instance of Closed Syllable Shortening (CSS), as illustrated in section 3.1. CSS applies in most dialects to turn a non-final CV:C syllable into a CVC syllable, as is shown in (48), with examples from Cairene and Ammani Arabic.

(48) *Cairene Arabic* (from Abu-Mansour 1992:49; Younes 1995:157):

k**aa**n 'he was'	~	k**u**nti 'you (f.sg.) were'
r**aa**H 'he went'	~	r**a**Hlaha 'he went to her'
binr**uu**H 'we are going'	~	mabinr**u**Hsh 'we are not going'
dur**uu**s 'lessons'	~	dur**u**shum 'their lessons'
Habb**ee**t 'I loved'	~	Habb**i**tha 'I loved her'

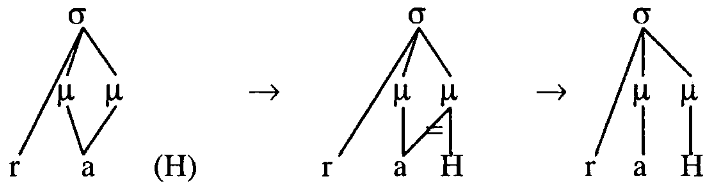
Ammani Arabic:

k**aa**n 'he was'	~	k**u**nti 'you (f.sg.) were'
r**aa**H 'he went'	~	r**a**Hfiiha 'he went with it (f.)'
bash**uu**f 'I see'	~	bash**u**fbiihum 'I see with them'
'**aa**l 'he said'	~	'**a**lli 'he said to me'

As was mentioned in section 3.1, this process can be best explained by means of Broselow's (1992) Bimoraicity Constraint (also see 3.3.1.2): since superheavy CV:C syllables are disallowed, CSS intervenes to shorten their vowels and make them bimoraic. The process applies especially when a consonant-initial suffix or clitic is added to a V:C-final stem (cf. the examples in (48) above): the last consonant, having lost its peripherality, is no longer extrasyllabic, and must therefore be syllabified in the preceding syllable, taking up the mora previously occupied by the second segment of the vowel (49).

(49)

a. b. c.



As was mentioned above, the implementation of CSS varies from one Arabic dialect to another. In most dialects, including AA, its domain of application appears to be the PW.[30] As can be seen from the examples in (50), the long vowel of the stem of a hollow verb like *shaaf* 'to see' undergoes shortening when followed by the subject suffix *–na* 'we', with which it forms a PW, but remains long when followed by the object suffix *–na* 'us' or by the noun *Nadhiir* (person's name), each forming a separate PW.

(50) (a) shuuf + na (subj.) → sh**u**fna 'we saw/ looked'

shuuf + na (obj.) → sh**uu**fna 'see/ look at us!'

shuuf + Nadhiir → sh**uu**f Nadhiir 'see/ look at Nadhiir!'

In Cairene Arabic, on the other hand, CSS also applies across PW-boundaries, within the PPh (Watson 2002:66). In this dialect, therefore, a PPh like *shuuf Nadhiir* surfaces as [ʃ**u**fnaðiːr], a CG like *shuufna* 'see/ look at us!' as [ʃ**u**fna].

As can be seen from the examples in (48), even in dialects that preserve long vowels in closed syllables before object suffixes, CSS does apply when a prepositional suffix is added. Some more examples from AA and Makkan Arabic are given in (51).

(51) *Ammani Arabic:*

'aal 'he said'	~	'**a**llu 'he said to him'
'eesh 'what'	~	'eshbiiha 'what (is wrong) with her?'
binjiib	~	binj**i**bfiiha

[30] Also compare Watson (2000:14, 2002:67-70) for San9ani Arabic.

'we bring' 'we bring in it (f.)'

Makkan Arabic (from Abu-Mansour 1992:passim):

siib 'leave!'	~	siblahum 'leave for them!'
ruuH 'go!'	~	ruHbaha 'go with it (f.)'
miin 'who'	~	minlaha 'who (will be) for her?'

Abu-Mansour (1992) argues that, while Bimoraicity remains a universal constraint on syllable well-formedness, the interdialectal divergences are due to the organization of the morpho-phonological component. According to him (ibid.:57), the peculiar implementation of CSS in Makkan Arabic (and in AA) derives from the fact that different suffixes belong to different morphological levels: agreement suffixes belong to level 1, prepositional clitics to level 2, object (and possessive) pronouns to level 3 (see 1.2). In these dialects CSS is an early principle, which stops operating after level 2, and is therefore restricted to words with level 1 (52b) and 2 (52c) affixes, thus differing from Egyptian Arabic.

(52)
(a) jaab 'he brought'

(b) jibna 'we brought'[31] (agreement suffix, level 1)

(c) jablaha 'he brought for her' (prepositional clitic, level 2)[32]

(d) jaabaha (*jabha) 'he brought it (f.)' (object pronoun, level 3)

(ibid.: passim)

[31] The alternation in the quality of the vowel is due to complex morpho-lexical processes, which are not of concern here.

[32] In AA, the long vowel is preserved when the clitic *-l-* 'for, to' is affixed, because the epenthetic vowel is inserted before the /l/, and not after, thus opening the preceding syllable (*jaa.bil.ha* 'he brought for her', rather than *jabliha*).

In a Prosodic Phonology framework, the issue can be analyzed in terms of phonological constituents: while level 1 and level 2 suffixes form a single PW with the stem, level 3 suffixes form independent PWs, thus behaving as clitics. Since in dialects like AA, Makkan and San9ani the domain of application of CSS is the PW, level 3 affixation cannot trigger the process.

Even in such varieties, however, syntactically linked words may in some cases form an accent group with its main stress on the second element, within which CSS takes place across PW boundaries (Watson 1999a:507; 2000:10). Examples from San9ani and AA are given in (53)

(53) *San9ani Arabic* (from Watson 2000:10, 14; 2002:68)

gaam + raHlih → gam raHlih 'he got up and went'

Ammani Arabic

wayn/ ween + raayiH → wen raayiH 'where are you going?'

kayf/ keef + kaanat → kef kaanat 'how was it (f.)?'[33]

In such instances, words like *wayn* 'where' and *kayf* 'how' are reduced to the status of proclitics to the following verbs (cf. Watson 2002:68).

Such accent groups, however, must be regarded as precompiled in the lexicon, since in dialects like AA and San9ani CSS does not seem to apply postlexically. In Cairene, on the other hand, CSS must apply postlexically, too (Abu-Mansour 1992:67-8; 72-3). This allows it to interact with high vowel deletion (see 4.3), as shown in (54).

(54) yaaxud 'he takes'	~	taxdi (< taaxudi) 'you (f.sg.) take'
SaaHib 'friend (m.)'	~	SaHba (< SaaHiba) 'friend (f.)'

(Abu-Mansour 1992:49-50)

[33] It is irrelevant whether the short vowel in these examples is derived directly from a diphthong or from a long vowel through monophthongization, since diphthong shortening can be considered an instance of CSS. In non-monophthongizing dialects like San9ani, in fact, the two processes apply exactly in the same contexts (Watson 2000:15-6; 2002:68).

maasik	~	maskiin (< maasik + iin)
'holding (m.sg.)'		'holding (m.pl.)'

(Watson 1999b:3)

In other Arabic dialects, among which AA, CSS appears to have a more limited application. Contrary to Cairene, in AA syncope does not trigger CSS (cf. AA *SaaHib* 'friend (m. sg.)' ~ *SaaHba* 'friend (f. sg.)', and not *SaHba* as in Cairene, AA *maasik* + *iin* > *maaskiin*, and not *maskiin* as in Cairene). Following Abu-Mansour (1992:73) in his observations about comparable forms in Makkan Arabic, we can thus conclude that CSS does not apply postlexically in AA.

Abu-Mansour (ibid.:65-6) also points out that Makkan Arabic stem-final closed syllables in bi-consonantal[34] and lame[35] verbs do not shorten even when followed by the prepositional clitics (55).

(55) ramee + t + l + ha → rameetalha (*rametlaha) 'I threw to her'

lammee + t + l + ha → lammeetalha (*lammetlaha) 'I gathered for her'

(ibid.:65)

The author (p. 67) explains these apparent exceptions on the basis of the consideration that the long vowel in these verbs is not underlyingly such, but derives from the diphthong /ai̯/ by means of monophthongization (see 2.3): when this latter applies, Makkan Arabic early CSS has already stopped operating.

In AA, however, forms with [e:] deriving from /ai̯/ do undergo shortening, as is illustrated in the examples in (56).

(56) ramayt + bi + ha → rametbiiha 'I threw with it (f.)'

[34] A bi-consonantal verb is one that has a geminated second radical (e.g.*Habb* 'to like, love', *HaTT* 'to put', *laff* 'to wrap, turn').

[35] Lame verbs (also called 'final weak verbs') are those biliteral verbs that end in a vowel and are diachronically derived from forms with a glide as a third radical, such as *rama* 'to throw' (< /r-m-j/) and *nisi* 'to forget' (< /n-s-j/).

lammayt + fi + ha → lammetfiiha 'I gathered in it (f.)'

wayn + rayiH → wen rayiH 'where (are you (m.sg.)) going?'

There is no need, however, to order monophthongization earlier in the derivation, since, as was pointed out above, CSS does not affect long vowels only, but also diphthongs. If it can be proved that diphthong shortening does not apply in Makkan Arabic, then the difference between the two dialects boils down to whether or not CSS affects diphthongs.

Another peculiarity of Makkan Arabic is that, though it does not exhibit CSS before object and possessive suffixes, it adopts another device to eliminate trimoraic syllables. This is the epenthesis of a vowel after the CV:C sequence, which opens the target syllable and makes it bimoraic (see examples in (57)).[36]

(57)

duruus 'lessons'	~	duruu.**sa**.hum 'their lessons'
nisiit 'I forgot'	~	nisii.**ta**.ha 'I forgot her'
feen 'where'	~	fee.**na**.ha 'where (is) she?'

(Abu-Mansour 1992: passim)

In other eastern dialects, however, neither CSS nor epenthesis apply before pronominal suffixes, leaving the CV:C syllable intact, as in AA (and generally Levantine) *kitaab* + *na* > *ki.taab.na* 'our book', Sudanese and Iraqi *beet* + *na* > *beet.na* 'our house' (cf. Broselow 1992:12-3, Watson 1999b:2).[37] In San9ani Arabic, *kitaab* + *naa* is realized alternatively as *ki.taab.naa* or *ki.taa.ba.naa*, depending especially on speech rate (Watson 1999a:513;1999b:3; 2000:16).

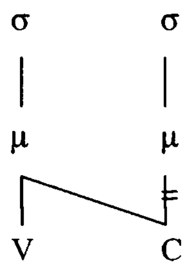
The presence of word-medial CV:C (i.e. trimoraic) syllables in dialects like AA, Sudanese, Lebanese, Iraqi, Gulf Arabic etc. clearly constitutes a challenge to the

[36] For a more accurate account of epenthesis in Arabic, see section 5.3.

[37] According to Abu Salim (1982, in Younes 1995), however, Syrian and Palestinian Arabic behave like Cairene in that their long vowels undergo CSS in closed syllables before any suffix, including pronominal clitics, although in these varieties CSS does not apply postlexically.

Bimoraicity Constraint. In order to preserve the bimoraic limit as a general feature of all Arabic dialects, and based especially on the distributional asymmetry between CV:C and CVCC syllables, Broselow (1992:14-19) proposes that the former be still considered bimoraic, i.e. 'derived from an adjunction rule that creates moras dominating two segments' (58).

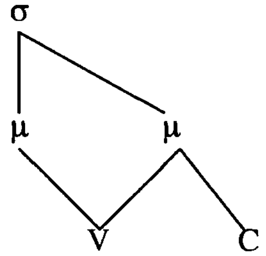
(58) Adjunction to Mora



(ibid.:14)

As was mentioned above, once material has been suffixed, the stem-final consonant can no longer be extraprosodic, because it is now in a non-peripheral position: it is therefore left stray. An analogous result is reached as a consequence of Vowel Deletion, when the syllable to which the relevant consonant belonged has lost its nuclear element and can no longer exist. The stray consonant is thus rescued by assignment to the preceding syllable, where it can share a mora with the long vowel (59).

(59)



(ibid.:15)

Laboratory work carried out by Broselow et al. (1995 and 1997) shows that the vowel of word-internal CV:C syllables in the relevant dialects is considerably shorter than that of CVV syllables but longer than that of CV syllables. This - in conjunction with data from Hindi and Malayalam[38] - appears to support the Adjunction-to-Mora hypothesis.

The ways in which the different dialects deal with PPh-internal CV:C syllables – i.e. preserving them, as in Sudanese, Levantine and others, shortening the vowel, as in Cairene, inserting an epenthetic vowel, as in Makkan – can be accounted for rather smoothly in an OT framework. As was mentioned above, the constraint *3μ disallows trimoraic syllables, and is therefore the main trigger of CSS. MAX-μ-IO militates instead against it, since it forbids the deletion of a mora. DEP-IO, which requires all segments in the output to have correspondents in the input, is also relevant to these data, since it rules out epenthesis. Finally, the constraint MAX-C-IO must be ranked high in all dialects, since none of them deletes the coda consonant of the CV:C syllable in order to satisfy *3μ.

In dialects such as Levantine, Iraqi and Sudanese, *3μ is ranked below MAX-μ-IO and DEP-IO, so that the super-heavy syllable is maintained. The tableau for AA *shuufna* 'see/ look at us!' is given in (60).

(60)

Input: /ʃu:f+na/	MAX-C-IO	MAX-μ-IO	DEP-IO	*3μ
a. ☞ ʃu:f.na				*
b. ʃuf.na		*!		
c. ʃu:.fa.na			*!	
d. ʃu:.na	*!			

In Cairene Arabic, where CSS takes place, it is MAX-μ-IO that is ranked lowest (see tableau in (61)).

[38] According to Broselow et al. (1997), in Hindi coda consonants are always assigned a mora (which yields a three-way distinction between monomoraic, bimoraic and trimoraic syllables), while in Malayalam coda consonants are always weightless, i.e. they always share a mora with the nuclear vowel. In languages like Levantine Arabic, 'coda consonants may be either weight-bearing or weightless, depending on the context: codas that follow a short vowel occupy their own mora, while codas following a long vowel share the preceding vocalic mora.' (ibid.:56).

(61)

Input: /ʃu:f+na/	MAX-C-IO	*3μ	DEP-IO	MAX-μ-IO
a. ʃu:f.na		*!		
b. ☞ ʃuf.na				*
c. ʃu:.fa.na			*!	
d. ʃu:.na	*!			

Finally, in dialects such as Makkan Arabic, DEP-IO must be ranked lowest in order to obtain vowel epenthesis (62).

(62)

Input: /ʃu:f+na/	MAX-C-IO	*3μ	MAX-μ-IO	DEP-IO
a. ʃu:f.na		*!		
b. ʃuf.na			*!	
c. ☞ ʃu:.fa.na				*
d. ʃu:.na	*!			

The account that has just been given does without the Adjunction-to-Mora option, by allowing *3μ to be ranked low. Broselow et al. (1995:123-126; 1997:65), on the other hand, assume that such a constraint is virtually unviolated in all dialects,[39] and thus what takes place in the Levantine-type dialects is not the surfacing of a trimoraic syllable, but Adjunction to Mora. A constraint forbidding mora sharing is thus postulated, which is called NOBRANCHINGMORA (NBM) in Broselow et al. (1995), NOSHAREDMORA in Broselow et al. (1997). In Levantine-type dialects, given the high ranking of *3μ, trimoraic syllables are ruled out just like in all other varieties; assuming that what takes place is instead Adjunction to Mora, in these dialects it is NBM that is ranked lowest (cf. tableau in (63)).

(63)

Input: /ʃu:f+na/	MAX-C-IO	*3μ	MAX-μ-IO	DEP-IO	NBM
a. ʃu:f.na		*!			
b. ʃuf.na			*!		
c. ʃu:.fa.na				*!	
d. ʃu:.na	*!				
e. ☞ ʃu·f.na[40]					*

[39] For many authors working within the moraic framework – among whom McCarthy and Prince (1986), Steriade (1991), Broselow (1992), the Bimoraic Constraint is in fact absolute.

[40] In this tableau, a half-long vowel is one that has undergone Adjunction to Mora.

In an OT framework refusing a stratal organization of the lexicon, however, it is difficult to explain why, within a given dialect, CSS applies before certain suffixes and not before certain others. Alignment and anchoring constraints are of no help here, because the affected segment is not located at the edge of any constituent.[41] The only possible solution is perhaps resorting to a compromise between Lexical Phonology and Morphology and OT, i.e. to Kiparsky's (2000) LPM-OT. This model adopts output-oriented constraints but accepts the existence of different morpho-phonological levels (namely, a stem, a word and a postlexical level), each of which is characterized by a specific constraint ranking. Also, the output of each level functions as an input to the successive level.

Abu-Mansour's (1992) – derived from Brame's (1974, see 4.3 and 5.3) – insight that the different behaviour of subject and object suffixes with respect to CSS is due to their belonging to different lexical strata is thus maintained: while subject suffixes are attached at the stem level,[42] object clitics are attached at the word level (Kiparsky 2000:356-9). We could thus assume that, at the stem level, *3μ (or NBM) dominates MAX-μ-IO, thus triggering CSS (64a), while at the word level the reverse ranking applies, yielding trimoraic syllables (or Adjunction to Mora) (64b).

(64) (a) *shuufna* [ʃufna] 'we saw/ looked'

Stem Level	*3μ (or NBM)	MAX-μ-IO
Input: /ʃu:f+na/		
a. ʃu:f.na	*!	
b. ☞ ʃuf.na		*

[41] It would actually be possible to obtain the correct outputs by means of high ranking of the constraint ALIGN-L (σ, PW). If *sh[u]fna* '(we) saw' is seen as made up of one PW and *sh[u:]fna* 'see/ look at us!' of two, this constraint triggers shortening in the former, but not in the latter item. However, as will be illustrated in chapters 4 and 5, ALIGN-L (σ, PW) cannot be active (i.e. it cannot outrank ALIGN-R (σ, PW)) in all of the dialects involved.

(b) *shuufna* [ʃu:fna] 'see/ look at us!'

Word Level	MAX-μ-IO	*3μ (or NBM)
Input: /ʃu:f+na/		
a. ☞ ʃu:f.na		*
b. ʃuf.na	*!	

The introduction of different morpho-phonological levels, however, makes of this model a constraint-based DT, rather than a version of OT.

3.3.2.1.2 Open Syllable Shortening

In dialects like Cairene and Palestinian, vowels in open syllables can also be shortened, as becomes apparent by comparing the two dialects with the more conservative MSA.

(65)

MSA	*Cairene*	*Palestinian*	
baa9u	**baa**9u	**baa**9u	'they sold'
baa9uuha	**ba**9uuha	**ba**9uuha	'they sold it (f.)'
	maba9uhaash	maba**(a)**9uhaash	'they did not sell it (f.)'
baa9atha	**ba**9itha	**baa**9atha	'she sold it (f.)'

(Younes 1995:158)

The Bimoraicity Constraint cannot be responsible for this phenomenon, which undoes a bimoraic syllable (CVV) to make it monomoraic (CV): higher prosodic structure must be involved. Younes (1995), who calls the process Open Syllable Shortening (OSS), takes into consideration stress placement, a factor which, as was mentioned above, is closely related to vowel quantity. According to this criterion, in Cairene OSS is a straightforward operation: 'a long vowel is shortened when it loses main stress due to the addition of stress-attracting suffixes' (ibid.:159). Younes thus concentrates on Palestinian Arabic (PA), where the implementation of OSS seems to have more complex implications, as shown in (66).

[42] Clearly, Abu-Mansour's (1992) level 2 affixes - such as the prepositional clitics - will also have to

(66) (a) OSS applies:

baabéen → **ba**béen — 'two doors'
'ariibáat → '**ari**báat — 'near (f.pl.)'
mabSuuTíin → mab**Su**Tíin — 'happy (m.pl.)'
makaatíbna → ma**ka**tíbna — 'our offices'
madaaríshum → ma**da**ríshum — 'their schools'
SaaHíbhum → **Sa**Híbhum — 'their friend'
shaafúuha → sh**a**fúuha — 'they saw her'
mashafuuháash → mashaf**u**háash — 'they did not see her'

(b) OSS does not apply:

Suurtéen (/Suurateen/) — 'two pictures'
T**aa**b9éen (/Taabi9een/) — 'two stamps'
sh**aa**Tríin (/shaaTiriin/) — 'smart (m.pl.)'
sh**aa**Tráat (/shaaTiraat/) — 'smart (f.pl.)'
9**aa**laméen — 'two worlds'
H**aa**rábato — 'she fought him'

Abu-Salim (1982, in Younes 1995:160-1) first put forward a metrical analysis of OSS in PA, stating that only vowels dominated by weak nodes are shortened. Younes (1995:162), however, points out that, in PA, long vowels in open syllables are shortened in only one specific environment, namely when unstressed and immediately followed by a stressed syllable. In the rural Palestinian dialect (RPA) analyzed by Younes (1995), however, one extra condition must be stated for OSS to apply, i.e. that the vowel in the following syllable be long besides being stressed (ibid.:164).[43] In the examples in (67), in which the stressed vowels are short, OSS is in fact blocked.

---

attach at the stem level.

[43] Similarly, Cantineau (1960:96) points out that, in Maghrebian Arabic, an unstressed long vowel is shortened when the word also contains a stressed long vowel (e.g. *kanúun* 'portable stove', *firáan*

(67) OSS does not apply:

mak**aa**tíbna 'our offices'
S**aa**Híbkum 'your (pl.) friend'
mad**aa**ríshum 'their school'

(ibid.)

Concluding that it is the length of the following vowel and not the weight of the following syllable that makes the difference, Younes (ibid.) formulates OSS as follows:

(68) VV → V / ___ $ CVV[44]

He (ibid.:165) further remarks that the fact that OSS also takes place before the feminine form of the *nisba* suffix (see (69)) is due to its *ijj* being underlyingly equivalent to *iij*.

(69) shaami 'from Damascus (m.sg.)' ~ sh**a**miyyi 'from Damascus (f.sg.)'
nuuri 'gypsy (m.sg.)' ~ n**u**riyyi 'gypsy (f.sg.)'
9iid 'the Eid, feast' ~ 9**i**diyyi 'for the Eid'

(ibid.:165)

This is in accordance with one of the tenets of this thesis, by which [ijj] and [i:j] would have a common UR X X X .
I

The author concludes that the domain of application of OSS in RPA is a disyllabic foot, and the process itself operates to maximize iambic feet (ibid.:169). In fact,

'mouse'). Also, as was mentioned above, some dialects exhibit a complete ban on words containing more than one long vowel.

[44] Younes (1995:165-6) shows that OSS also applies before exceptional CVCCC syllables, and he proposes that the few real exceptions to this process be justified on functional grounds.

following McCarthy and Prince (1990), and contradicting what was said in section 1.4, he maintains that feet in Arabic are right headed.[45]

As for AA, it seems to behave in a similar fashion to PA with respect to OSS. For comparison's sake, I have reported in (70) the same lexical items as in the PA examples in (66).

(70) (a) OSS applies:

'ariibáat → 'a**ri**báat 'near (f.pl.)'

mashafuuháash → mashaf**uh**áash 'they did not see her'

mabSuuTíin → mab**Su**Tíin 'happy (m.pl.)'

makaatíbna → mak**at**íbna 'our offices'

madaaríshum → mad**ar**íshum 'their schools'

(b) OSS does not apply:

shaafúuha → sh**aa**fúuha 'they saw her'

baabéen → **baa**béen 'two doors'

Suurtéen (/Suurateen/) → **Suu**rtéen 'two pictures'

shaaTríin (/shaaTiriin/) → sh**aa**Tríin 'smart (m.pl.)'

shaaTráat (/shaaTiraat/) → sh**aa**Tráat 'smart (f.pl.)'

Taab9éen (/Taabi9een/) → T**aa**b9éen 'two stamps'

9aaladméen → 9**aa**laméen 'two worlds'

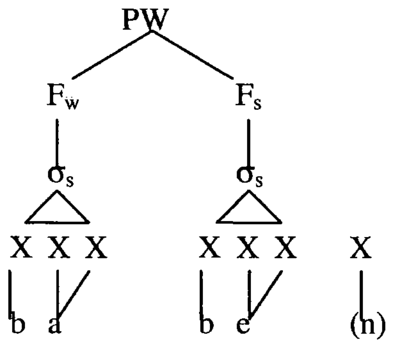
SaaHíbhum → **Sa(a)**Híbhum 'their friend'

Haarábatu → **Ha(a)**rábatu 'she fought him'

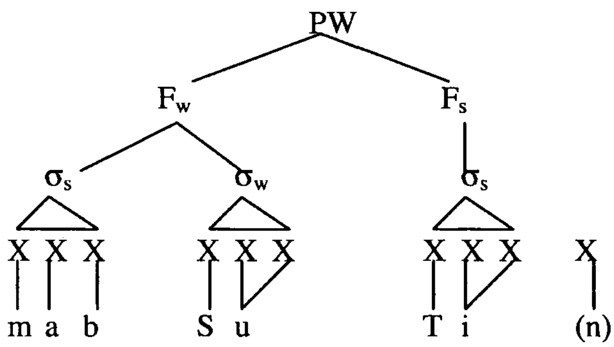
[45] Watson (2002:87) actually argues that, even though most Arabic dialects – and especially those of the Mashreq – exhibit trochaic systems, 'much of Arabic morphology [...] is based on the optimum iamb: a light syllable followed by a heavy syllable'.

As in PA, and contrary to RPA, in AA unstressed vowels in open syllables seem to undergo OSS before any stressed syllable, whether or not this contains a long vowel. There are, however, some clear exceptions to this pattern. In particular, when the target vowel is word-initial - as in *baabéen* or *shaafúuha* – OSS does not apply. The different behaviour of PA and AA in these cases is likely to be related to foot structure: if we consider AA feet to be trochaic rather than iambic, then a word-initial syllable will be the head of a foot, and will thus be assigned secondary stress (primary stress is assigned to the rightmost foot in the word).

(71) (a) *bàabéen* 'two doors'



(b) *màbSuuTíin* 'happy (m.pl.)'

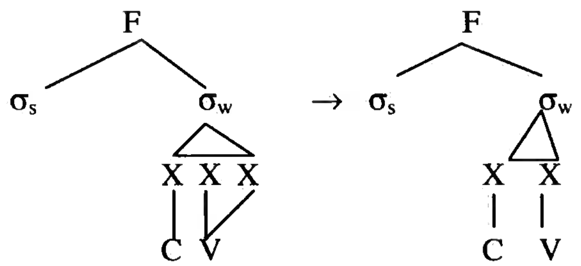


As was mentioned in 3.1, a vowel bearing prominence (71a) is less prone to shorten than an unaccented vowel (71b). If feet were iambic, as Younes (1995) claims them to be in RPA, then the initial syllable in *baabeen* would belong to a foot whose head is the syllable [be:n], and would thus be unstressed. Clearly, this would not explain its resistance to shortening in AA. The same reasoning applies to the other non-shortened forms, such as shaaT(i)ráat 'smart (f.pl.)' or 9aalaméen 'two worlds'.

In the case of *baabeen* and *shaafuuha*, foot isochrony – i.e. the tendency of vowels in monosyllabic feet to be longer than vowels in polysyllabic feet – works in the same direction, since the first syllable of these words occupies a foot of its own. In other words, if the first /a:/ in *bàabéen* were shortened, an ill-formed monomoraic foot would result.[46]

In conclusion, AA long vowels undergo OSS when they occupy a weak position within the foot. This process can thus be formulated as follows:

(72) OSS in AA:



From an OT perspective, OSS can be accounted for if the constraint WSP (heavy syllables are stressed) is made to refer to secondary stress as well as main word stress. Its effect will thus be to ban heavy syllables - and consequently long vowels - in completely unstressed positions. The tableau in (73) shows that, by ranking WSP above MAX-μ-IO, the optimal candidate for *mabSuuTiin* - i.e. the one with a shortened second vowel – is selected. WSP must also dominate EVEN-TROCHEE, since OSS produces the uneven foot $[mab.Su]_F$.

(73)

Input: /mabSU:T+I:n/	WSP	MAX-μ-IO	EVEN-TROCHEE
a. $[map.Su:]_F$ $[Ti:n]_F$	*!		
b. ☞ $[map.Su]_F$ $[Ti:n]_F$		*	*

[46] These considerations, however, do not explain the optional shortening that words like *Sa(a)Híbhum* 'their friend' and *Ha(a)rábatu* 'she fought him' undergo. The alternations may perhaps indicate an on-going change affecting this dialect.

The tableau in (74) shows instead how, in the case of the non-shortened *9aalameen*[47], WSP is irrelevant and the optimal candidate is thus selected by MAX-μ-IO. In order to obtain the correct result, this latter constraint must outrank EVEN-TROCHEE.

(74)

Input: /9a:lam+aIn/	WSP	MAX-μ-IO	EVEN-TROCHEE
a. ☞ [9a:.la]F [me:n]F			*
b. [9ala]F [me:n]F		*!	

3.3.2.1.3 Final Vowel Shortening

Pre-pausal vowels (and consonants) shorten in MSA, as in *'adduwalu lkubraa* > *'adduwalu lkubra* 'the superpowers' (Holes 1995:50). In many dialects, PW-final long vowels are only preserved when they are morphologized and maintain grammatical distinctions, as in Gulf Arabic *smá9u* 'they heard' vs. *sma9úu* 'they heard him' or Palestinian and Syrian *bíddi* 'I want' vs. *biddíi* 'I want it (m.)' (Holes 1995:54). Cantineau (1960:95) already notes a similar tendency in Classical Arabic, where word-final long vowels appear to often shorten, especially in poetry, as in *'anaa* > *ana* 'I', *-naa* > *-na* (1st pl. pronoun suffix), *l-'aydii* > *l-'aydi* 'the hands', *nabghii* > *nabghi* 'we desire'.

Examples from AA of Final Vowel Shortening (FVS) are given in (75).

(75)	sama9 + u:	→	sama9u	'they heard'
	sama9 + ni:	→	sama9ni	'he heard me'
	walad + i:	→	waladi	'my boy/ son'
	wajad + na:	→	wajadna	'we found/ he found us'
	bidd + i:	→	biddi	'I want'
	bidd + u:	→	biddu	'he wants'
	ghani:	→	ghani	'rich (m.sg.)'
	9aadi:	→	9aadi	'ordinary (m.sg.)'
	xalli:	→	xalli	'allow!'

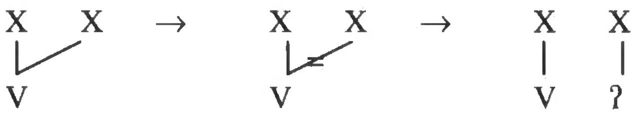
[47] *9aalameen* has been preferred over *baabeen* as an example because, in this latter case, the effect of further constraints on foot structure (e.g. FT-BIN) would have made the contrast with *mabSuuTiin* less

Watson (1999a; 2000; 2002) assumes that, in dialects such as Cairene Arabic, FVS is a diachronic process. However, at least as far as AA is concerned, consistent alternations prove that the process is instead synchronic. As the examples in (76) show, whenever a suffix is added to the above-listed forms their last vocoid, having lost its peripheral position, is realized as long.

(76)	xall[i] 'allow! (sg.)'	~	xall[i:]ni 'allow (sg.) me!'
	wajadn[a] 'we found'	~	wajadn[a:]ki 'we found you (f.sg.)'
	bidd[i] 'I want'	~	bidd[ij]ak 'I want you (m.sg.)'
	9aad[i] 'ordinary (m.sg.)'	~	9aad[ijj]a 'ordinary (f.sg.)'

Abumdas (1985:46) remarks that in Libyan Arabic, when FVS applies, a weak [ʔ] is often created. This shows that the slot (or mora) left empty by the second segment of the vowel is not always deleted. When it is not, a default segment is inserted to fill in the empty position (77).

(77)



Interestingly, this seems to be a sort of reverse process to Compensatory Lengthening, though Watson (1989) would rather talk of 'diphthongization' (cf. 2.3 for discussion).

In OT, the constraint that triggers FVS is *FINAL-VOC:, requiring word-final vocoids to be short. As the tableau of *ghanii* 'rich (m.sg.)' in (78) shows, this constraint has to be ranked above MAX-μ-IO for the shortening to take place.

(78)

Input: /ɣanI:/	*FINAL-VOC:	MAX-μ-IO
a. ɣani:	*!	
b. ɣani		*

clear.

3.3.2.2 Lengthenings

Confirming a cross-linguistic tendency (Kager 1999:5) that may translate into a constraint against long vowels (see 3.1), in Arabic lengthening processes are much less common than shortenings. The few instances that can be found are either sporadic and idiosyncratic in character, or have applied diachronically and are now fossilized, or exhibit strong morpho-lexical restrictions. Therefore, they will be dealt with very briefly.

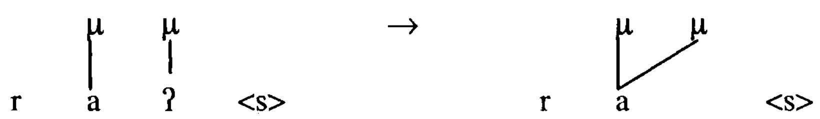
3.3.2.2.1 PU-final lengthening

According to Watson (1989:157; 1999a:502; 2000:5), in San9ani Arabic short vowels do not occur in PU-final position, which seems to indicate that a process of domain-final lengthening takes place in that dialect. In AA, the phenomenon is sporadic and difficult to formulate in phonological terms, probably due to the conflicting effects of FVS, which does not seem to apply in San9ani.

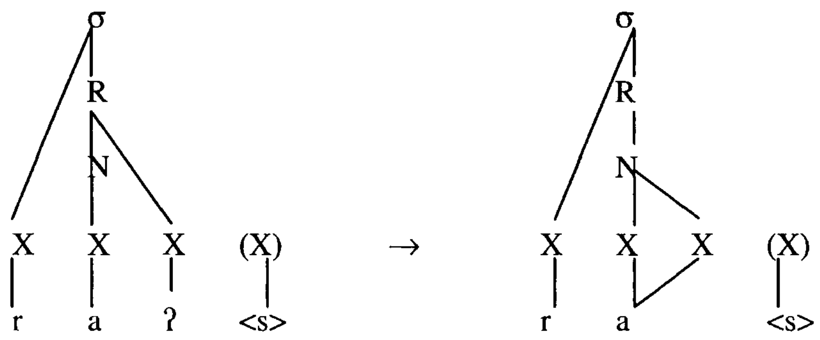
3.3.2.2.2 Compensatory Lengthening

In AA, as in most Arabic dialects, Compensatory Lengthening (CL) is found especially as a diachronic phenomenon. In varieties such as AA and Cairene, a word like Classical Arabic *ra'ʔs* 'head' has become *raas*, first undergoing pre-consonantal ʔ-deletion, and then lengthening of the /a/, which took up the slot left empty by the deleted segment. Watson (2000:4) argues that this is most economically represented in terms of moraic theory, where the mora left floating by the deletion simply relinks to the vowel, without any further operation (79a). In an X-slot framework, the X freed by the deletion has to disassociate from the rhyme node and to reassociate with the nucleus node (79b).

(79) (a)



(b)



3.3.2.2.3 Particle lengthening

In AA, as in other dialects (for Libyan Arabic, see Abumdas 1985:245-6, for Cairene Arabic, see Broselow 1992 and Watson 1999a), the last vowel of a particle which is not already long is lengthened before a consonant-initial pronominal suffix,[48] as in *li* + *kum* > *liikum* 'for you (pl.)', *bi* + *hum* > *biihum* 'with them (pl.)'. A particle-final consonant, too, is lengthened before a pronominal suffix, as in *min* + *ka* > *minnak* 'from you (m. sg.)', *9an* + *hu* > *9annuh* 'about him'.

Since this phenomenon is restricted to a handful of lexical items, and both the triggering and the affected words belong to specific morphological categories, we can assume that these forms have become fossilized and are now precompiled in the lexicon (cf. Hayes 1990).

[48] When a vowel-initial suffix is added, a long vowel may result from assimilation, as in *bi* + *ak* > *biik* 'with you (m. sg.)'.

CHAPTER 4: VOWEL DELETIONS

4.1 Elision, Syncope, Truncation

This section portrays the phenomenon of Vowel Deletion (VD) in its different forms, and is divided into three subsections. The first of these deals with deletion as a reduction phenomenon and gives an account of the various ways of representing it. The second part analyzes in more detail the different types of VD (namely elision, syncope and truncation), their relation with prosodic structure and the factors conditioning their application. The last subsection introduces the OT constraints and constraint rankings that may account for VD in its various forms.

4.1.1 Deletion as reduction

The deletion of a segment is a process by which such a segment, though underlyingly present, is not phonetically realized in certain contexts. It is a very common phenomenon cross-linguistically, although it takes on different forms, and can be seen as an extreme case of reduction. Diachronically, in fact, a segment often undergoes different stages of reduction before being deleted; in the dialect of Genoa, for instance, Latin/SI intervocalic [t] has first lenited to [d] and then to [ð], before deleting completely (cf. SI [venu:to] 'come (past part. m.sg.)' vs. Genoan [veɲyu]). In other Northern Italian varieties, moreover, the process has not gone beyond the lenition stage.

Donegan (1985:39) actually considers reduction processes as a subgroup of lenition processes, together with assimilations. She (ibid.:37) stresses the fact that lenition processes generally implement the principle of least effort, by maximizing ease of articulation.

> Unlike fortitions, lenition processes function to produce more-optimal sequences of segments; they make sequences more pronounceable by assimilating the properties of one segment to those of a neighboring segment, by deleting segments, and by

> substituting segments that are 'weaker' in some respect for those that are 'stronger'.[1]
>
> (ibid.:38)

Lenitions and fortitions, however, may often be regarded as opposite sides of the same process. This can certainly be said of deletions and insertions, which, as will be shown in more detail in ch. 5, could be dealt with as an alternation [segment] ~ [0]. Surely at the postlexical level, where already formed words are inserted into higher prosodic structure, it makes more sense to talk about the deletion of a pre-existing sound. As will be shown, however, deletions are often morphologically or lexically conditioned, and this is when a [segment] ~ [0] alternation could be a more useful representational tool.

The deleted segment can either be syllabic or non-syllabic, a vocoid or a contoid. For the purposes of this thesis, however, only Vowel Deletion (VD) will be dealt with in some depth.

The process being a weakening one, short vowels are more prone to undergo deletion than long vowels, although the latter can also be deleted. This happens if the loss of material affects the melodic rather than the skeletal tier, or if the phonotactic constraints of the language do not allow a less radical solution (see Yawelmani in section 4.1.2.1).

As will be shown in more detail below, VD is also connected to stress placement, in that an unstressed vowel is more likely to be deleted than a stressed one. Donegan (1985:38) actually notes that lenition processes in general tend to affect segments in 'weak' positions, such as 'unstressed vowels, syllable-final consonants, segments in intonational 'valleys', etc.'. Moreover, the 'intrinsic' quality of a vocoid heavily affects its propensity towards deletion. Less sonorous vowels are generally more prone to undergo deletion or other weakening processes than more sonorous ones (compare GF in ch. 2). In several Japanese dialects, for example, vowel height conditions the application of weakening processes such as vowel devoicing, VD or

[1] Notice that the fact that lenited segments may be 'weaker' or 'more easily pronounceable' does not necessarily imply their being less marked. A segment like [d], for instance, is cross-linguistically more common than its lenited counterpart [ð].

tone lowering. These apply only to the high vocoids /I/ and /U/, and not to the non-high ones. Haraguchi (1984) shows that this is strictly connected to the inverse relation between the height and sonority of a vocalic segment (cf. 1.1). That is, the notion of height is only indirectly relevant to the application of weakening processes, through the notion of sonority. Tone lowering, destressing, deletion, devoicing, reduction are all forms of lenition that are enhanced by a lower degree of sonority. Conversely, non-high vocoids, being more sonorous, are more prone to undergo strengthening processes or to resist weakening. In Haraguchi's (ibid.:155) words, 'the least sonorous vowels are more likely to undergo devoicing and deletion unless they are protected from these 'weakening' processes by stress, tone, tension, certain vowel qualities, or surrounding sonorous consonants'.

In Japanese, for instance, a morpheme-final high vowel belonging to the first member of a compound which is a loanword from Chinese is deleted when surrounded by voiceless consonants, as in (1).

(1)	koku	koku-go	kok-kyoo
	'country'	'mother tongue'	'border, boundary'

(ibid.:145)

In the Kagoshima dialect, moreover, high vowels are systematically deleted in all morpheme-final positions, as in (2).

(2)	UR	HVD	PF
	tuki 'moon'	→ tuk	→ tuʔ

(ibid.:147)

It should be noted, in connection with what was said above, that morpheme-final vowels are located in what Donegan (1985) designates as 'a weak position'. In fact, as will be shown below, these cases cannot be attributed to clash-avoiding nor to quantity sensitive VD.

The implementation and extension of lenition processes is often affected by speech rate, so that deletions and reductions apply more frequently and more extensively in fast and casual speech. As Donegan (1985:39) puts it, 'the domain of application of lenition processes increases when articulatory effort is likely to be low and when informal situations or highly predictable content make lower perceptual qualities acceptable'.

4.1.1.1 Representation of Vowel Deletion

In linear phonology (Chomsky and Halle 1968), the structural change associated with VD is represented as in (3).

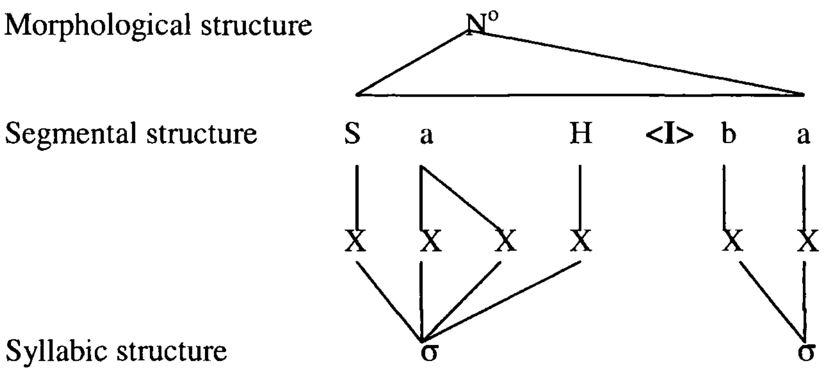
(3) V → 0

This kind of representation, however, fails to show that VD, along with epenthesis, is strictly connected with syllable structure (Lapointe and Feinstein 1982, Archangeli 1984, Noske 1993, Broselow 1992, etc.). Most deletion processes are syllable-sensitive, i.e. they apply in order to bring the syllabic structure of a string closer to the 'prototypical' – i.e. the least marked, universally or language-specifically – syllable. This is, in general, CV, CV: or CVC. Moreover, all VD processes actually affect syllable structure, even those that do not seem to be affected by it. This is due to the simple fact that vowels are syllable peaks, and the only obligatory elements in a syllable: their presence or absence is bound to modify the syllabic asset of a string.

In an autosegmental model, VD often takes the form of a process by which a vocalic segment is left unattached to either the skeleton or the syllabic tier. According to Lapointe and Feinstein's (1982:80) 'dual representation' model, for instance, a deleted segment is attached to a morphological node but not to a syllabic node. A word like Arabic *SaaHiba* [Sa:Hba] 'friend (f.)', whose /i/ has undergone deletion, would thus be represented as in (4).

(4)

UR: /Sa:HIba/



PF: [Sa:Hba]

(adapted from Lapointe and Feinstein 1982)

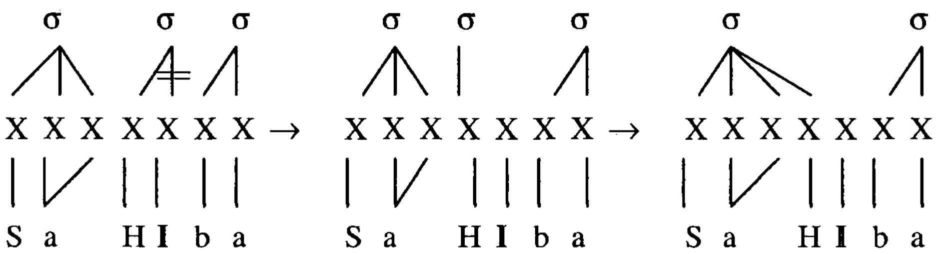
As far as the relation between segmental and timing structure is concerned, there actually seem to be 3 different types of deletion:

1. On the skeleton, where a timing slot is lost and the corresponding segment is therefore left floating, and may be retrieved in case an empty timing slot and the corresponding subsyllabic node are made available (see Burzio's (1987) analysis of prosodic reduction in Italian, in 4.2).
2. On the melodic tier, in which case a timing slot on the skeleton is left unfilled, thus triggering temporal compensation - by which other melodic material goes to fill in the empty slot (cf. Compensatory Lengthening in ch. 3) - or even some sort of epenthesis.
3. On both melodic tier and skeleton, in which case the lost segment can never be recovered (see Burzio 1987:14).

Traditionally, VD is described as the action of a phonological process eliding a vowel or at least a vowel matrix unspecified for any features inside an already built prosodic structure. Thus, the assumption is that it takes place after initial syllabification.

According to this interpretation, the deleted segment is delinked from the syllable node and left floating, as in (5). As a consequence, it cannot be phonetically realized.

(5)



As is shown in (5), the immediate output of VD contains vowel-less syllables, which in many languages are ill-formed. Therefore, some syllable structure readjustment must take place. In this case, the vowel-less syllable is deleted and its onset consonant is resyllabified into the coda of the preceding syllable.

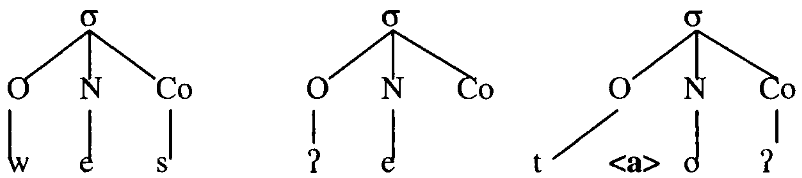
This type of syllable restructuring, however, is no longer necessary if a different approach to this phenomenon is adopted. Many phonologists, among whom Lapointe and Feinstein (1982), Itô (1989), Broselow (1992), Noske (1993) and Majdi and Winston (1994), have pointed out that VD, just like Vowel Insertion (cf. ch. 5), is a prosodic phenomenon which is a function of the universal processes which relate segmental strings with syllable structures. In this regard, Broselow (1992:32) remarks that

> Many vowel/ zero alternations may result not from the effect of a rule applying to a phonological representation that is already prosodically organized, but rather in the course of the association of consonantal melody to prosodic template.

In order to avoid arbitrary resyllabification, Lapointe and Feinstein (1982:72) hypothesize that processes involving the creation or deletion of syllables can only operate as part of a universal syllabification algorithm, applying before syllable-sensitive processes. According to these authors (ibid.:78), however, 'the Syllable Assignment Algorithm (SAA) [...] can apply each time a word-level structure is generated by the morphology', so that a form of cyclic resyllabification is postulated.

Other scholars adopt instead a templatic approach to syllabification, by which a segmental string is mapped onto a series of syllables characterized by a given template (cf. Itô (1989), Noske (1993) in section 1.3). In this framework, a deleted segment is one that cannot be linked to a subsyllabic node in the first place, due to a mismatch between the CV sequence and the syllabic template. Specifically, VD takes place when a vocoid meets an obligatory marginal position that it cannot fill, and is thus skipped by the syllabification mechanism. Since such a segment is left floating, it cannot be phonetically realized. A string like Tonkawa /we+s?eta+o?/ 'he cuts them' will thus be syllabified as in (6). The direction of syllabification being right to left, the vowel /a/ is matched to an onset position that would require a lower sonority segment, and is therefore left unparsed.

(6)



However, if a language exhibits lexical syllabification, a post-lexical VD process will still require delinking and relinking of segmental material to syllabic structure.

4.1.2 Types of Vowel Deletion

The two main factors triggering VD are both related to syllable structure. The first one is vocoid clash avoidance, required in order to satisfy the same OCP constraint against strings of consecutive syllable peaks that operates in the case of GF (cf. 3.1). The second factor is to do with syllable minimality: a language-specific constraint against subminimal (generally monomoraic) syllables requires that these are deleted wherever possible. Other types of deletion, however, seem to be syllable-insensitive. One is deletion by weak position, e.g. morpheme-final truncation.[2] Another type is deletion by assimilation, an example of which is loss of a vocoid between voiceless

2 As was mentioned in 3.2, truncation may often be construed as triggered by foot structure, rather than by syllable structure, constraints.

consonants, when deletion becomes an extreme case of devoicing (see Japanese in (1) above).

4.1.2.1 Vowel Deletion as a clash-avoiding device

When two vowels – i.e. two syllable peaks - come into contact, a clash takes place. As was explained in ch. 2, there appears to be a universal constraint against such clashes, whose strength varies cross-linguistically. In some languages, there is an absolute ban on vowel clashes, while in other languages such clashes are avoided only in so far as this does not involve the violation of some other, stronger constraints. The ban against vowel clashes is closely related to the Onset Principle, in that a language that does not allow onsetless syllables will obviously lack uninterrupted strings of syllable peaks.

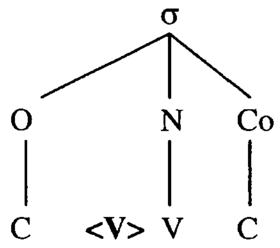
The process by which a constituent-final vowel is lost when followed by another vowel is traditionally called *elision*, and is a special type of deletion. The reverse process, by which a vowel at the beginning of a constituent is lost when preceded by a vowel at the end of the preceding constituent is traditionally called *prodelision*. In this thesis, however, the more common term Vowel Elision (VE) will be used to refer to both processes.

In traditional linear phonology, VE was represented in its basic form as in (7).

(7) $V \rightarrow 0 / \_\_\_\_ V$

As was mentioned in 4.1.1.1, however, this representation is clearly insufficient for the purpose of showing the real nature of the process. The main point here is not the deletion of a segment, but its being unattached to the syllabic tier, and consequently remaining phonetically unrealized. An unrealized vowel can be schematically represented as in (8).

(8)



In a templatic approach, since the syllabification process consists of the matching of a series of syllabic matrices to the CV string, a mismatch takes place whenever a given segment is incompatible with the next subsyllabic constituent available. In this case, the mismatch consists of a vocoid being paired up with an O position. When, because of the high sonority of the vocoid or of language-specific phonotactic constraints, a glide cannot be formed, the vocoid is left floating, and cannot be phonetically realized.

By the same token, when two identical vowels clash, rather than forming an on-glide + vowel sequence (in those languages where these are allowed) or merging into a long vowel, one of them may be elided, whatever its sonority.

As was said above, however, sonority is often relevant to the application of clash avoiding processes when the two clashing vowels differ in quality. When a more sonorous vocoid clashes with a less sonorous one, it is often the latter that undergoes elision. For instance, Mirror Image Deletion in Greek – as described by NV (159-60) and Nespor (1993:200) - is a clash-avoiding process by which of two consecutive vowels the weaker in sonority (according to a sonority scale /i/ < /e/ < /a/ < /o/) is deleted, as shown in (9).

(9)

[se aɣa'po]	→	[saɣa'po] '(I) love you'
[ta 'eperne]	→	['taperne] '(he) took them'

(from Nespor 1993:ibid.)

It should be remembered, however, that GF may also apply as a result of a clash between vowels of different sonority, as an alternative to VD. In the languages in which GF overrides VD, elision seems to apply more readily to lower, more sonorous vowels (see Italian in 2.2 and 4.2).

The distribution of stress, too, may play a role in determining which vocoid will undergo deletion when a clash occurs. Most often, however, prominence on the target vocoid blocks the process altogether (see Italian in 4.2). As was pointed out in 4.1.1, stressed segments are stronger and therefore less prone to undergo deletion than unstressed ones.

When sonority and stress placement are in conflict – i.e. when a more sonorous unstressed vocoid clashes with a less sonorous stressed vocoid - one of two things appears to happen. Either elision does not apply, and both vocoids (syllable peaks) are preserved, or one of the two factors prevails. If stress prevails, the unstressed vowel is deleted, even if it is more sonorous. If sonority prevails, the least sonorous vocoid is deleted, even if it is stressed.

The above mentioned Greek process of Mirror Image Deletion is an example of the latter type. Stress placement does not affect it: an accented vowel, if weaker in sonority, can be deleted, and the stress is transferred onto the remaining vowel, as in (10).

(10)

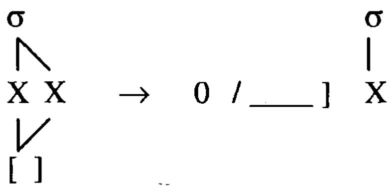
[ta 'exo]	→	['taxo] '(I) have them'
[to 'ipa]	→	['topa] '(I) said it'

(from NV: 159; Nespor 1993:200)

However, the most significant factor in determining which vowel will be elided when a clash occurs appears to be the position of the vocoids with respect to each other: in some languages, the leftmost clashing vowel is invariably deleted, in other languages, the rightmost one.

In Yawelmani, a much studied dialect of the Amerindian language Yokuts (Kuroda 1967, Archangeli 1984, Noske 1993), it is always the first vowel that is deleted, even when it is long. This is the result of the fact that, in this language, both hiatuses and diphthongs are banned (Noske 1993:107), so that GF from the second segment of a long vowel is not an available option even when such a vocoid is high. Archangeli (1984:196) formulates the rule of VE in Yawelmani as in (11), where [ ] is an underspecified segment and X is a V slot on the skeleton.

(11) *Vowel Elision*[3]



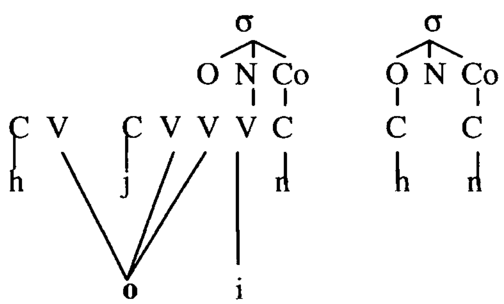
Noske (1993:107), however, argues that, if the process is considered to be the result of syllable structure constraints, rather than deriving syllable structure from the process itself, the excessive complexity of rules like the one in (11) can be avoided. In the framework he adopts, as was mentioned above, VD is interpreted according to a templatic model where all subsyllabic positions are projected for all syllables. When a vowel is immediately preceded by another vowel, there is no consonant to be linked to the projected onset position, which is obligatorily filled in Yawelmani. Since, according to Noske (ibid.), in this dialect syllabification applies from right to left, when the mechanism reaches the leftmost clashing vowel the onset of the syllable to its right is not yet filled; as a consequence, there is a mismatch between syllabic and skeletal structure and the linking is blocked. Therefore, as an emergency measure, the syllabification mechanism skips the vowel and goes on to match the onset position to the next skeletal element to its left. When the vowel is long, its first V-slot is also ignored. As was mentioned above, the vowels that are thus skipped by syllabification

[3] Archangeli (1984) only refers to long vowels because, according to her, in Yawelmani there are no morphemes ending in a short vowel.

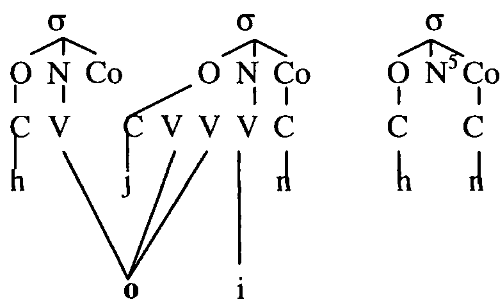
are not phonetically realized. The process is exemplified in (12) on the word [hojinhin] < /hojo: + in + hn/ 'was named'.

(12) UR: /hojo: + in + hn/[4]

(a) The right-to-left parse reaches the long vowel, which cannot be matched with the next available obligatory position, namely O:



(b) The parsing mechanism skips the mismatched V-slots, which remain floating:

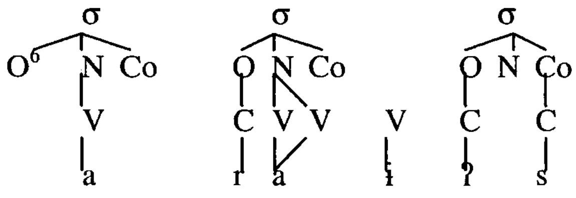


In Tigrinya, contrary to Yawelmani, it is the rightmost vowel that is deleted when a clash takes place. According to Noske (ibid.), this is due to the fact that directionality of syllabification is left to right in this language. Moreover, the preceding vowel has to be long, because Tigrinya allows two short vocoids in the same rhyme, forming a long vowel or a diphthong (Noske 1993:116). Left-to-right syllabification and VE in Tigrinya are exemplified in (13).

[4] Like Arabic, Yawelmani exhibits a non-concatenative morphology, by which consonants and vowels are located on different tiers.

[5] This empty nuclear position will be filled by an epenthetic vowel.

(13) /ara:iʔs/ 'heads'

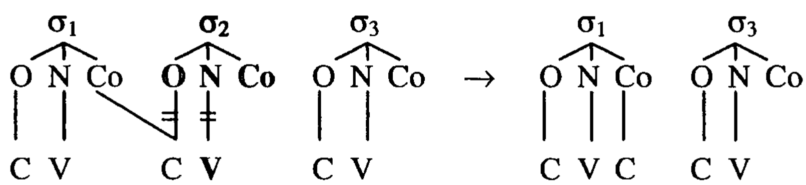


Noske (ibid.) points out that, since onsets must be obligatorily filled in this language, when syllabifying a post-vocalic consonant C_1 the mechanism has to look ahead to the next segment. If this is a consonant, C_1 can be syllabified in the coda of the preceding syllable; if it is a vowel, C_1 must be syllabified in the onset of the following syllable. He (ibid.) thus concludes that right-to-left directionality is the unmarked case, this being probably related to the Maximal Onset Principle (cf. 1.3).

4.1.2.2 Vowel deletion as related to syllable quantity

As was mentioned above, another main factor triggering VD is a cross-linguistically common constraint against monomoraic syllables, or at least against sequences of such syllables. When a similar configuration occurs, various strategies are adopted by languages to repair it. As was said in ch. 3, one such strategy is the lengthening of the nuclear vocoid of the target syllable, in order to make the latter bimoraic (OSL). Otherwise, a syllable node may be deleted and its onset consonants resyllabified into the coda of the preceding syllable, whenever this is compatible with the phonotactics of the language. The vocoid occupying the peak of the deleted syllable is therefore left stranded and, once again, cannot be realized phonetically. The corresponding timing slot is also lost.

(14)

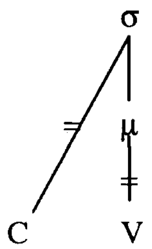


[6] Both this onset and the last nucleus will be filled in by epenthetic segments.

The diagram in (14) shows schematically how a subminimal CV syllable (σ_2 here) is deleted, leaving its two segments floating. While the onset consonant may be rescued by the coda of the preceding syllable (σ_1 here), the nuclear vowel is lost.

Since the process is clearly to do with syllable quantity, it can easily be accounted for in terms of moras. As Broselow (1992:32) points out, in some languages syncope seems to get rid of monomoraic syllables wherever possible. It may therefore be represented as in (15), which reads 'delete a syllable which contains only one mora'.

(15) Syncope:



Lapointe and Feinstein (1982:89) point out that '[...] constraints on sequences of syllable types all follow from the metrical structure constraints of the language'. As was mentioned above, in their model VD takes place as part of a Syllable Assignment Algorithm (SAA): because universal and language-specific constraints rule out certain candidate syllable structures for a word, this is forced to assign trees with unattached (and hence 'deleted') vowels.

Stress placement plays a fundamental role here, too, since a constraint often demands that a syllable undergoing deletion be unstressed (but see counterexamples from Greek in (10)). For cases in which an unstressed vowel in the middle of a word is deleted, the term *syncope* has traditionally been used. In English, for instance, unstressed vocoids – and therefore unstressed syllables – often undergo syncope, as in *memory* ['mɛmɹi] (alternation *mem[0]ry* ~ *mem[ɔ:]rial*) or *secretary* ['sɛkɹətɹi] (alternation *secret[0]ry* ~ *secret[æ]rial*). Hammond (1997:47) points out that, in American English, an unstressed vowel is syncopated in the following three cases:

a. If it is word-initial; or

b. If it immediately follows the only stress in the word and it is not in the last syllable of the word; or

c. If it is one of two stressless syllables intervening between two stressed syllables.

Based on these remarks, Hammond (ibid.) concludes that American English syncope is not only a matter of syllable structure, but also of foot structure: monomoraic syllables are deleted in order to obtain optimal, binary feet (see 4.1.3). For instance, syncope of the second vowel in a word like *ópera* yields the disyllabic foot ['ɑ:pɹə], syncope of the first vowel in *paráde* yields the completely footed, bimoraic ['pɹeɪ̯d].

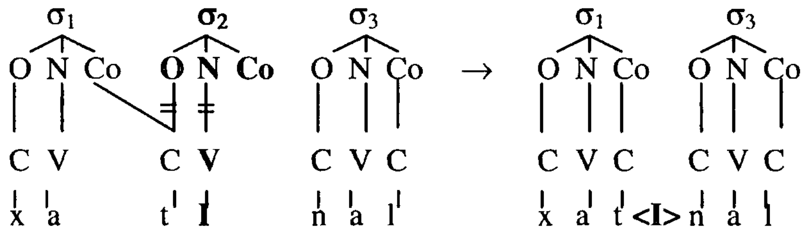
Though not to such an extent as stress assignment, the sonority of the target vowel may also affect the application of syncope. Clearly, less sonorous vowels are more prone to undergo the process. For example, as will be shown in 4.3, in many Arabic dialects only high vowels, which are less sonorous, are deleted in light syllables.

In languages such as English or Arabic, syncope is a postlexical phenomenon: it is blind to grammatical structure, exceptionless, optional and productive. Other languages, instead, exhibit lexical syncope.

In Yawelmani, for instance, syncope applies to some specific morphemes to delete a vowel in a monomoraic open syllable which is preceded by another open syllable, as in (16).[7]

[7] Noske (1993:92) actually questions the existence of what Kuroda (1967) calls 'Two Sided Open Syllable Deletion'. Such a process, in fact, only seems to undo the results of Epenthesis in the majority of cases (see 5.1). Moreover, Kuroda (1967) appears to assume such a process 'chiefly in order to

(15) /xat-In+al/ 'eat (mediopassive dubitative)'



(adapted from Lapointe and Feinstein 1982:87-8)

In Klamath, too, VD in monomoraic syllables is a lexical process. It applies to the first short vowel in a prefixed stem, provided the relevant syllable is open (see (17)).

(17) $[paga]_{st}$ 'smokes' BUT $[sna]_{af}[pga]_{st}$ 'cause to smoke'

(from Lapointe and Feinstein 1982:96)

When syncope applies to a string of the form CVCV, the issue arises of how to determine which of the two vowels will be deleted. This varies from language to language and even from variety to variety within the same language (see Arabic in 4.3), so that the choice seems to be based on parameter setting. In a templatic approach, whether the leftmost or the rightmost vowel is deleted depends on the direction of syllabification and footing (see 4.3 for further discussion).

4.1.2.3 Truncation

Some VD processes, however, do not appear to be triggered by syllable structure constraints, even if such constraints still play a role in determining when they should be blocked. One such process is so-called *truncation*, i.e. the deletion of an unstressed vowel in word-final position. In Yawelmani, for instance, Verb Final Vowel Deletion

---

explain the behaviour of one single affix: mediopassive *–in-*' (ibid.:93). Rather than a lexical process

(VFVD) takes place, by which the final vowel of a verb ending is deleted when the preceding stem terminates in an open syllable, as shown in (18).[8]

(18) /taxa:/ 'bring' + /k'a/ (imperative) → [taxak'] 'bring it'

(from Noske 1993:91)

This type of process seems to be motivated mainly by the fact that, as was mentioned in 4.1.1, morpheme-final segments are in one of those 'weak positions' that tend to favour reduction of the selfsame segment (see Donegan 1985).

Although syllable structure can hardly be considered a trigger for Yawelmani VFVD, it must be pointed out that the process is blocked precisely when it would result in ill formed syllable structure. If it took place after closed syllables, then it would create complex codas, which are disallowed in this language.

4.1.3 Vowel Deletion in OT

Deletions have been treated in different ways in the various subtheories of OT. PS's Containment Theory, assuming that the input form is entirely contained in the output postulates that no input element is actually removed, even when it is not pronounced: deleted segments are still present in the output, only they are left unparsed, and consequently unrealized. Therefore, the constraint PARSE(-SEG), which requires underlying segments to be parsed into syllable structure, counters deletion.

In Correspondence Theory (McCarthy and Prince 1995), the constraint that militates against the deletion of an input segment is MAX-IO, which belongs to the MAXIMALITY family. It enforces segment preservation by stating that 'input segments must have output correspondents' (Kager 1999:67). More specifically, VD violates

proper, therefore, this may be a case of phrasal phonology precompiled in the lexicon (see Hayes 1990).
[8] Also see Italian in 4.2.

the constraint MAX-V-IO, which requires faithfulness to input vowels. Therefore, for VD to apply, MAX-V-IO must be dominated by one or more markedness constraints.

As was specified above, VD may be triggered by different factors, i.e., in an OT perspective, by different types of markedness constraints.

Since VE originates in a vowel clash, OT will clearly characterize its application as due to high ranking of the constraint NOHIATUS (or ONSET). When this latter dominates MAX-V-IO – and so do DEP-C-IO and *M/V, precluding the alternative solutions of Consonant Insertion and GF respectively – one of the clashing vowels is elided. Tableau (19) shows that, given this ranking, the optimal candidate is the one exhibiting VE.

(19)

Input: /CVVC/	NOHIATUS	DEP-C-IO	*M/V	MAX-V-IO
a. CV.VC	*!			
b. CV.[C]VC		*!		
c. CGVC			*!	
d. CVC				*

Vowel syncope, on the other hand, may be seen as a device to optimize foot structure and reduce the number of unparsed syllables in a string. Foot optimization constraints such as FT-BIN, when ranked above PARSE-SYL, may yield unparsed syllables. However, if PARSE-SYL dominates MAX-V-IO, the extra syllables can be eliminated by deleting their vowels.[9] The candidates for the parsing of a hypothetical string CVCVCV are evaluated in tableau (20). MAX-V-IO being ranked lowest, the optimal candidate is the one with the syncopated vowel.

[9] Clearly, a more direct trigger of syncope is $*\sigma_{\mu}$, the constraint disallowing monomoraic syllables (see ch. 3). Also, since syncope is often limited to light syllables which are adjacent to other light syllables, a constraint *Light Light has also been proposed (Zawaydeh 1997:201).

(20)

Input: /CVCVCV/	FT-BIN	PARSE-SYL	MAX-V-IO
a. (CV.CV.CV)	*!		
b. (CV.CV)CV		*!	
c. ☞ (CVC.CV)			*

By the same token, syncope may be triggered by alignment constraints such as ALL-FT-L/R - by which every foot stands at the left/right edge of the PW - dominating PARSE-SYL, in its turn dominating MAX-V-IO. Because of ALL-FT-L/R being highly ranked, the requirement that all syllables should be parsed cannot be implemented by simply parsing syllables into more than one foot per word (iterative footing),[10] otherwise some such feet will necessarily fail to be aligned with the PW at the relevant edge. However, PARSE-SYL can still be at least partially implemented at the expense of the lower ranked MAX-V-IO: by deleting a vowel, a potential syllable will also be eliminated, thus trivially satisfying PARSE-SYL. The interaction of these constraints is shown in (21), in relation to the string CVCVCVCV.

(21)

Input: /CVCVCVCV/	ALL-FT-L[11]	PARSE-SYL	MAX-V-IO
a. (CV.CV)(CV.CV)	*!		
b. (CV.CV)CV.CV		**!	
c. ☞ (CV.CVC).CV		*	*

Avoidance of marked foot types such as (LL), i.e. high ranking of RH-CONTOUR (see 3.1), may also trigger vowel syncope. If both this constraint and DEP-μ-IO - by which an underlyingly short vowel cannot be lengthened to form a canonical iamb (LH) - dominate MAX-V-IO, an input vowel is deleted in order to allow the other light syllable to become heavy and thus form a less marked foot ([CV.CV] → [CVC]), as in the tableau in (22).

[10] Kager (1999:179) points out that iterative footing would be the result of a different ranking of the three constraints, namely PARSE-SYL, MAX-V-IO » ALL-FT-X.

(22)

Input: /CVCV/	RH-CONTOUR	DEP-μ-IO	MAX-V-IO
a. ☞ (CVC)			*
b. (CV.CV)	*!		
c. (CV.CV:)		*!	

The same applies to trochaic feet, which are optimally even: if RH-CONTOUR (EVEN-TROCHEE) and MAX-μ-IO (the constraint against vowel shortening) outrank MAX-V-IO, a string CV:CV is restructured as CV:C by means of syncope of the last vowel, as shown in (23).

(23)

Input: /CV:CV/	RH-CONTOUR	MAX-μ-IO	MAX-V-IO
a. ☞ (CV:C)			*
b. (CV.CV)		*!	
c. (CV:.CV)	*!		

Faithfulness to consonants may block VD in order for consonants to be syllabified, when MAX-C-IO and, for example, *COMPLEX are ranked high with respect to PARSE-SYL, even if this latter dominates MAX-V-IO. That is, if consonants cannot form complex onsets and/ or codas and cannot be deleted, then a vowel is retained in order to constitute the nucleus of a syllable in which the consonants can be syllabified, even if such a syllable may not be parsed, thus violating PARSE-SYL. The ranking of $*\text{COMPLEX}^{\text{ONS}}$ with respect to $*\text{COMPLEX}^{\text{COD}}$ may determine the position of the deleted vowel in cases in which two (or more) different vowels in a word are possible targets.

Alignment constraints have also been proposed in order to account for the different locations of syncope (see Abu Mansour 1995). In particular, ALIGN-L (σ, PW) and ALIGN-R (σ, PW) require every syllable to be aligned with, respectively, the left or right edge of some PW. These constraints are gradient, and violations are measured by the number of moras that separate each misaligned syllable from the relevant edge. Such a number changes according to the way in which segments are syllabified, which

[11] If ALL-FT-R were used instead of ALL-FT-L, the tableau would be very similar.

in its turn may depend on the location of syncope. In a word of the form /CV:CV$_1$CV$_2$CVC/, for instance, Align-L (σ, PW) will receive less serious violations if V_2 is deleted than if V_1 is, while Align-R (σ, PW) is violated less seriously by deletion of V_1. The tableau in (24) shows how, in a language with the ranking Align-L (σ, PW) » Align-R (σ, PW), the optimal candidate is the one in which V_2 is syncopated.[12]

(24)

Input: /CV:CV$_1$CV$_2$CVC/	*Light Light	Align-L (σ, PW)	Align-R (σ, PW)
a. CV:.CV$_1$.CV$_2$.CVC	*!	μμ μμμ μμμμ	μμ μμμ μμμμ
b. CV:C.<V_1>CV$_2$.CVC		μμμ μμμμ!	μμ μμμ
c. ☞ CV:.CV$_1$C.<V_2>CVC		μμ μμμμ	μμ μμμμ

Vice-versa, in a language with the ranking Align-R (σ, PW) » Align-L (σ, PW) the winning candidate will exhibit syncope of V_1.[13]

The deletion of a particular vowel is often rendered by means of constraints such as No[i] or No[ə] (see Kager 1999:283-4), stating that that vowel is not allowed in light syllables. In such constraints, however, reference to syllable structure is only implicit and has to be stipulated. It could be made explicit by reformulating them as No-i]$_\sigma$ or No-ə]$_\sigma$ or, better still, by ranking identity constraints with respect to $*\sigma_\mu$. For instance, the ranking Ident-IO[a] » $*\sigma_\mu$ » Ident-IO[i] calls for deletion of [i], but not of [a], in monomoraic syllables. This latter solution has the advantage of making use of independently motivated constraints.

The fact that VD is often blocked when the target vowel is stressed can be accounted for by means either of the constraint Max-V́-IO, requiring faithfulness to stressed vowels, or of the more indirect Pk-Max-IO, by which a segment which is the stress peak in the input must have a correspondent that is the stress peak in the output. The

[12] A candidate in which no syncope applies is eliminated by high ranking of some markedness constraint, here *Light Light (see fn.9 above). Moreover, since such a candidate has one more syllable than the two candidates that undergo syncope, it must also violate both syllable alignment constraints more seriously.

[13] Zawaydeh (1997), however, shows that syllable alignment constraints cannot account for differences in output for epenthesis and syncope in certain Arabic dialects. These issues will be dealt with in more detail in section 4.3 and ch. 5.

fact that segments which are stressed in the base are not deleted in affixed forms can be ascribed to high ranking of the OO-correspondence constraint HEADMAX-BA, which requires that every segment in the base's prosodic head - i.e. every segment that is stressed in the base - should have a correspondent in the affixed form (see 4.3).

The constraint family ANCHORING-IO penalizes the application of deletions and insertions at the edge of constituents (see 1.5). For instance, ANCHORING-IO(GW, R) requires that any segment at the right periphery of the output GW has a correspondent at the right periphery of the input GW. The contiguity constraints of the type CONTIG-IO, on the other hand, by requiring that input-output mappings involve contiguous substrings, militate against medial deletion or epenthesis. In particular, medial deletion is countered by I-CONTIG, which requires a certain domain to be a single contiguous string in the input (Kager 1999:250).

As for truncation, it can often be construed as the result of the action of foot well-formedness constraints such as RH-CONTOUR (cf. 3.2, 4.2). The relation between the truncated form and its base offers among the best examples of OO-interaction, since truncated forms often preserve phonological properties of their base even when these are not contextually motivated (see 1.5).

4.2 Vowel Deletion in Italian

This section deals with vowel elision, truncation and syncope in Italian. The first subsection analyzes Vowel Deletion (VD) when triggered by a vowel clash, i.e. Vowel Elision (VE). Two main VE processes can be found in this language: Lexical and Postlexical VE. In the second section, the complex truncation phenomena of Italian and their phonological, lexical and syntactic conditioning are illustrated. Only a brief paragraph is dedicated to syncope, because this process has little relevance in Italian. An OT analysis of the phenomena at hand is given whenever appropriate.

4.2.1 Vowel Deletion as related to clash avoidance

VD in Italian often applies as a clash-avoiding device, just as GF (see 2.2). It may actually take place when this latter process is blocked, e.g. when none of the clashing vowels is high, or when one of them is stressed.

While sonority – and therefore height – is crucial to GF, it does not seem to be relevant to VD in Italian. Only, less sonorous vocoids are more prone to undergo GF than VD, which instead is the only clash-avoiding device available in the case of non-high vowels. However, in a language like Italian the Onset Principle is not fully implemented, and onsetless syllables are therefore allowed, not only word-initially (e.g. *andare* [**an**.da:.re] 'to go', *orologio* [**o**.ro.lɔ:.dʒo]'clock, watch'), but also word-internally (e.g. *Paolo* [pa.**o**.lo] 'Paul', *aereo* [**a**.**ɛ**.re.**o**] 'aeroplane'). This means that the ban on sequences of two or more consecutive syllable peaks is not absolute, and clash-avoiding devices such as VD or GF (cf. ch. 2) may fail to apply.

In fact, VD applies obligatorily only within the PW, while across PWs it is an optional fast speech process, with some lexical exceptions that will be dealt with below.

4.2.1.1 Lexical Vowel Elision

VD inside the PW, which we may call Lexical Vowel Elision (LVE), is a morphologically conditioned operation, which deletes an unstressed vowel which is followed by another vowel only when a morpheme boundary separates the two segments (cf. (25)).

(25)	fam**a** + oso	→	famoso 'famous'
	giall**o** + astro	→	giallastro 'yellowish'
	Bar**i** + ese	→	barese 'from Bari'

BUT

paura[14] → *pura 'fear'

meandri → *mandri 'meanders'

(from NV:29)

The process only requires the two vowels to belong to separate morphemes, without differentiating between types of morpheme boundaries. For instance, VD applies both in (26a), which is an example of derivation, and (26b), which is an example of inflection.

(26) (a) [[bell**o**]$_A$ezza]$_N$ → [bellezza]$_N$ 'beauty'

(b) [[bell**o**]$_A$ i]$_A$ → [belli]$_A$ 'beautiful (m.pl.)'

(from Nespor 1993:181)

The deleted vowel is always the first one, i.e. the one belonging to the stem, whatever its sonority. If such a vowel is stressed, however, the process is blocked, as can be seen in (27).

(27) virt**ú** + oso → virt**u**oso (*virtoso) 'virtuous'

caff**é** + ina → caff**e**ina (*caffina) 'caffeine'

(from Nespor 1993:82)

[14] Note that in the word *paura* [pa'u:ra] GF does not apply, since the high vowel [u] is stressed (cf. 2.2).

The process is clearly a clash-avoiding device, since it only takes place when a vowel clash occurs. The stem-final vowel is not deleted when the following suffix begins with a consonant, as in (28).

(28)	appari + zione	→	apparizione 'apparition' (*apparzione)
	ama + bile	→	amabile 'amiable' (*ambile)

(from Nespor 1993:83)

Nespor (ibid.) notes that the reason why the stem-final vowel in a word like *amabile* is not deleted is not related to syllable structure constraints, since a word like **ambile* would be a perfectly well-formed Italian word (cf. *ambito* 'environment, sphere').

By looking at the examples in (29), however, one can conclude that the application of VD is restricted to cases of suffixation, while prefixed forms are not affected.

(29)	pre + avvíso	→	preavvíso 'notice' (*pravvíso)
	extra + uteríno	→	extrauteríno 'extra-uterine' (*extruteríno)
	pro + ávo	→	proávo 'great grandfather' (*právo)

(from Nespor 1993:182)

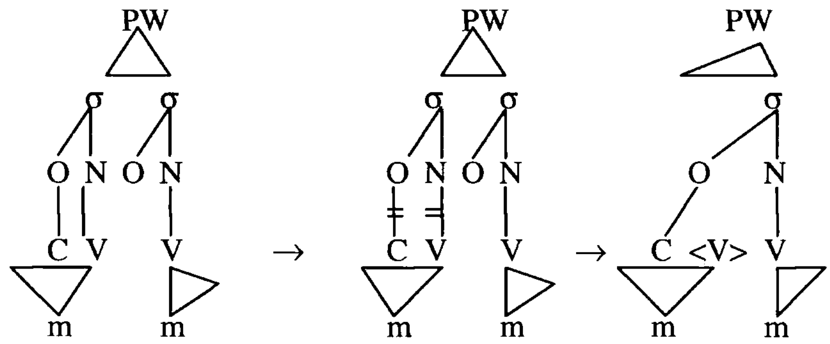
The few cases in which VD optionally (but by no means obligatorily) applies, such as *sovresposto/ sovraesposto* 'overexposed' or *seminterrato/ semiinterrato* 'basement; half-planted', can be seen as more lexicalized items. Since in this case the process is not productive, nor obligatory, it has to be considered as a diachronic phenomenon separate from the productive process that applies to suffixed forms (ibid.).

Since Italian prefixes, at least when they end in a vowel (see NV), do not form one PW with the stem they are attached to, while suffixes do (cf. 1.2), Nespor (1993) concludes that the reason why productive VD does not apply between a prefix and its host is that the domain of application of this process is the PW. VD can thus be construed as a PW-span operation, as in (30).

(30) V → 0 / [____ + V]$_{PW}$
[-high]

The diagram in (31), which includes only the most relevant prosodic and morphological information, makes the connection with syllabic structure more explicit.

(31)



This analysis is confirmed by the fact that in compounds, too, VD is neither obligatory nor productive;[15] in fact, the two members of a compound form two separate PWs in Italian.

The process applies cyclically, as can be seen in (32).

(32) tavolo + ino → tavolino 'small table'

tavolino + etto → tavolinetto 'very small table'

[15] Here again, exceptions like *terzultimo* (not **terzoultimo*) 'third from last' or *soprabito* (not **sopraabito*) 'overcoat' are due to a higher degree of lexicalization.

The morphological information which is required here is merely structural, and the operation applies to all morphologically complex words in which the environment is met. Other deletion processes of Italian, on the other hand, are not only structurally, but also lexically restricted.

In an OT framework, as was mentioned in 4.1, LVE is obtained by ranking NOHIATUS above MAX-V-IO. The tableau in (33) illustrates how the optimal candidate for *famoso* 'famous' is selected.

(33)

Input: /fama + oso/	NOHIATUS	MAX-V-IO
a. ☞ [fa'mo:zo]		*
b. [fama'o:zo]	*!	

For words like *caffeina* [kaffe'i:na] 'caffeine', it has been mentioned that what blocks LVE must be the presence of a stress on the base form *caffé* 'coffee'. Such a stress, however, is not present in the surface form of the derived word. In OT, therefore, this VD blocking must be construed as a case of OO-correspondence between base and affixed form. In particular, a constraint HEADMAX-BA states that each segment which is stressed in the base must have a correspondent in the affixed form; if this outranks NOHIATUS, deletion of a vowel which has a stressed correspondent in the base is blocked (see tableau (34)).

(34)

Input: /kaffɛ + Ina/ Base: [kaf'fɛ]	HEADMAX-BA	NOHIATUS	MAX-V-IO
a. [kaf'fi:na]	*!		*
b. ☞ [kaffɛ'i:na]		*	

In order to limit the application of the process to PW-internal environments, deletion must be prevented from applying at the right edge of PWs. This can be done by ranking the constraint ANCHORING-IO(PW, R) - declaring that an input segment at the right edge of a PW must have a correspondent in the output – above NOHIATUS, as shown in the tableau of the prefixed form *extrauterino* 'extra-uterine' (35).

(35)

Input: /ɛkstra # uterino/	ANCHORING-IO(PW, R)	NOHIATUS	MAX-V-IO
a. ☞ ɛkstra]$_{PW}$ [uterino		*	
b. ɛkstr] $_{PW}$ [uterino	*!		*

LVE, however, is not only blocked outside the PW, but also within morphemes. In accordance with this observation, it has often been noted that faithfulness tends to be stricter in non-derived environments, i.e. inside the root (Mascaró 1976, Kiparsky 1982a). In OT terms, root-faithfulness » faithfulness (Kager 1999:76). In this specific case, a constraint MAX-V-IO(root) must be ranked above NOHIATUS, while the more general MAX-V-IO must still be ranked below it. Tableau (36) shows how monomorphemic *meandro* 'meander' escapes LVE.

(36)

Input: /meandro/	ANCHORING-IO(PW, R)	MAX-V-IO(root)	NOHIATUS	MAX-V-IO
a. ☞ me.an.dro			*	
b. man.dro		*!		*

Since, in the relevant environments, LVE is always preferred over Consonant Insertion and GF, both DEP-C-IO and *M/V must outrank MAX-V-IO. Forms like *barese* [bare:ze] 'from Bari', in which a stem-final /I/ is deleted rather than being turned into a glide (cf. ill-formed *[barje:ze]), are evidence that not only *M/a, *M/e and *M/o, but also *M/I and *M/U – i.e. the constraints banning high vocoids from syllable margins – dominate MAX-V-IO. GF from high vowels, however, is preferred over VE in other environments, and namely across PW boundaries (cf. *Bari e Brindisi* [ba:rjebrindizi] 'Bari and Brindisi').[16] In 2.2, in fact, the ranking *M/a, *M/e, *M/o »

[16] It also applies in some PW-internal environments, like that of non-elided *virtuoso* [virtwo:zo] 'virtuous'. This case, however, is taken care of by the constraint HEADMAX-BA, which disallows VE without disallowing GF, since it refers to the realization of a segment and not to its syllabic status. The evaluation of the candidates for this word is illustrated in tableau (i).

MAX-V-IO » *M/I and *M/U was postulated for Italian. This ranking paradox cannot be solved in terms of traditional, monostratal OT, since the constraint banning VE across PW boundaries, ANCHORING-IO(PW, R), also bans GF. In DT, on the other hand, this issue is easily explained by classifying LVE as a lexical process, GF as a postlexical one. Quite naturally, the solution seems to lie in a compromise between Lexical Phonology and OT, i.e. in Kiparsky's (2000) LPM-OT. This constraint-based model, as was explained in 3.3, allows for the existence of different morpho-phonological strata, each having its own constraint ranking. The output of one stratum constitutes the input of the next one. In our case, it is sufficient to refer to a lexical and a postlexical level: while at the former *M/I and *M/U dominate MAX-V-IO, the opposite ranking order obtains at the latter. Besides, ANCHORING-IO(PW, R) must be ranked low in the postlexical stratum. The tableaux for *baresi al mare* [bare:zjalma:re] 'people from Bari at the seaside' in (37) show that lexical ranking yields LVE within the PW *baresi* [baresi], while postlexical ranking triggers GF at the juncture of the two PWs *baresi al* [bare:zjal].[17]

(i)

Input: /virtU+oso/ Base: [vir'tu]	HEADMAX-BA	ANCHORING-IO(PW, R)	NOHIATUS	DEP-C-IO	*M/U	MAX-V-IO
a. [vir.tu.o:.zo]			*!			
b. ☞ [vir.two:.zo]					*	
c. [vir.to:.zo]	*!					*
d. [vir.tu.Co:.zo]				*!		

[17] The constraints HEADMAX-BA and MAX-V-IO(root) have been left out because they are not relevant to the examples.

(37)

Lexical level	ANCHORING-IO(PW, R)	NOHIATUS	DEP-C-IO	*M/I	MAX-V-IO
Input: /barI+esI/					
a. [ba.ri.e.si]		*!			
b. [ba.rje.si]				*!	
c. ☞ [ba.re.si]					*
d. [ba.ri.Ce.si]			*!		
Postlexical level[18]	DEP-C-IO	NOHIATUS	MAX-V-IO	*M/I	ANCHORING-IO(PW, R)
Input: /ba.re.si#al#ma.re/					
a. [ba.re:.zi.al.ma:.re]		*!			
b. [ba.re:.zal.ma:.re]			*!		*
c. ☞ [ba.re:.zjal.ma:.re]				*	*
d. [ba.re:.zi.Cal.ma:.re]	*!				

4.2.1.2 Postlexical Vowel Elision

Despite what was said above, clash-avoiding VD may also apply across PW boundaries in Italian. This happens especially when no high vowel is involved in the clash, so that GF is blocked, although VE may also, occasionally, affect high vowels (see 2.2). Contrary to LVE, postlexical VE (henceforth PVE) is generally an optional process, enhanced by an increase in speech rate. Both (38a) and (38b) are acceptable in spoken Italian, and the faster the speech rate, the more likely the utterance is to correspond to (38b).

(38) (a) E' un uom**o** **a**llegro.

(b) E' un uom**[0]** **a**llegro.

'(He) is a cheerful man'

However, Nespor (1990:382-4; 1993:209) points out that, if the two adjacent vowels are identical, the second one is deleted, provided it is not stressed (cf. (39a) vs. (39b)). This process should be construed as an example of VD, or better, as Vowel Degemination (VDeg), rather than as a Vowel Shortening phenomenon, since the clashing vocoids belong to separate PWs. In (39c), stress on the first clashing vowel does not block VDeg, confirming that it is the rightmost vowel that is deleted.

(39) (a) Piantáva álberi. '(S/he) planted trees' [pjan.ta:.**va.al**.be.ri] (not *[pjan.ta.**val**.be.ri])

(b) Piantáva abéti '(S/he) planted fir trees' [pjan.ta:.**va**.be:.ti]

(c) Sará arriváto '(He) will have arrived' [sa.**rar**.ri.va:to]

When a stressed syllable immediately follows the rightmost clashing vowel, however, contraction (see ch. 3) replaces VDeg, as in (40).

(40) Pianterá abéti '(S/he) will plant fir trees' [pjan.te.**ra:**.be:.ti] (not *[pjan.te.**ra**.be:.ti])

This seems to be connected with foot minimality, since deletion of the second clashing vowel would yield a sub-minimal monomoraic foot, while contraction into a long vowel ensures the foot $[ra:]_F$ bimoraicity.

Contrary to VD outside the PW, VDeg is not a fast speech process, but a connected speech phenomenon,[19] whose domain of application can be identified with the PU

[18] Notice that Vowel Lengthening in stressed open syllables (see 3.2) and the sonorization of intervocalic /s/ are assumed to be the result of postlexical constraint interaction. The relevant constraints, however, are not shown, in order not to overcomplicate the tableau.

[19] Nespor (1990:383-4) shows that VDeg exhibits the typical features of prosodic rules: it may affect any of the X-bar categories, it applies across phonetically empty syntactic constituents, it is blocked by an intervening pause.

(Nespor 1990:383). This is exemplified in (41), where the process applies across an IPh (41a), but not across a PU boundary (41b).

(41) a. Robert**o**, **o**rmái non piú giovane, ci guardava dalla riva.

'Roberto, (who was) not young anymore, watched us from the shore' (> robert**o**rmai)

b. Ho mandato Robert**o**. **O**rmái dovrebbe averla trovata.

'(I) have sent Roberto. By now (he) should have found her' (*robert**o**rmai)[20]

Moreover, it should be noted that this process applies only at the juncture between two PWs, and not when the two vowels belong to the same PW, as in *aree* 'areas' (Nespor 1990:383; 1993:209). It is therefore a PW juncture operation within the PU.

In an OT perspective, the fact that, in Italian, a clashing non-high vocoid never forms a glide but may be elided can clearly be accounted for by ranking MAX-V-IO below the constraints banning non-high vocoids from marginal positions. MAX-V-IO also seems to be outranked by DEP-C-IO, since VE is preferred over Consonant Insertion. Free ranking must instead be postulated between MAX-V-IO and NOHIATUS: if the latter constraint dominates the former, PVE takes place, otherwise the hiatus is maintained. The tableaux in (42) report the evaluation of the candidates for *uomo allegro* 'cheerful man' according to the two possible subhierarchies.

(42)

(a) Subhierarchy 1 (NoHiatus » Max-V-IO)

Input: /wɔmo allegro/	*M/o	Dep-C-IO	NoHiatus	Max-V-IO
a. [wɔ:.mo̯al.le:.gro]	*!			
b. [wɔ:.mo.Cal.le:.gro]		*!		
c. [wɔ:.mo.al.le:.gro]			*!	
d. ☞ [wɔ:.mal.le:.gro]				*

(a) Subhierarchy 2 (Max-V-IO »NoHiatus)

Input: /wɔmo allegro/	*M/o	Dep-C-IO	Max-V-IO	NoHiatus
a. [wɔ:.mo̯al.le:.gro]	*!			
b. [wɔ:.mo.Cal.le:.gro]		*!		
c. ☞ [wɔ:.mo.al.le:.gro]				*
d. [wɔ:.mal.le:.gro]			*!	

In Italian, as was pointed out in 2.2, free ranking must also be postulated between *M/U, Max-V-IO and NoHiatus, since /U/ may sometimes be deleted when a clash occurs.

As in the case of GF, the fact that a stressed vocoid escapes VE and VDeg can be accounted for by ranking Pk-Max-IO above NoHiatus (see 4.1 ad ch. 2). The choice of the vowel to be deleted is instead problematic to explain in OT terms. Such an explanation should be couched in the ranking of Anchoring-IO(PW, R), by which any segment at the right edge of the output PW must have a correspondent at the right edge of the input PW, with respect to Anchoring-IO(PW, L), demanding that any segment at the left edge of the output PW has a correspondent at the left edge of the input PW. Since the affected vowel is the leftmost when the two clashing vocoids are different and the rightmost when they are identical, we are faced with a ranking paradox: in the first case, Anchoring-IO(PW, L) should dominate Anchoring-

[20] If the two PUs are construed as having a strong semantic link, however, they can be restructured into one PU, thus yielding the surface form [robɛrtormai].

IO(PW, R), while in the second case ANCHORING-IO(PW, R) should dominate ANCHORING-IO(PW, L). It must be noted, however, that an alternative DT analysis is also unsatisfactory, since it can only be based on stipulation.

As for examples like that in (40), where two identical clashing vocoids contract into a long vowel, they are straightforwardly accounted for by the action of the constraint FT-BIN, requiring that feet be binary under a syllabic or moraic analysis, and thus banning monomoraic feet.

4.2.1.2.1 Lexically conditioned VE

There is a restricted number of words which are lexically marked as far as VE is concerned. These include certain numerals (e.g. *vent'anni* 'twenty years' vs. *venti mesi* 'twenty months'), determiners (e.g. *quest'albero* 'this tree' vs. *questo cane* 'this dog') and verb forms (*dev'essere* 's/he must be' vs. *deve fare* 's/he must do') which undergo elision far more regularly than other items. This is shown by the fact that the deletion has become encoded in the spelling practice. For some of these words, however, elision of the final vowel is obligatory (cf. (43a)), while for all others it is optional (cf. (43b)).

(43) (a)l'albero 'the tree' (*lo albero)

un urto 'a crash' (*uno urto)

quell'orco 'that ogre' (*quello orco)

bell'amico 'nice friend' (*bello amico)

buon intingolo 'good sauce' (*buono intingolo)

nessun usignolo 'no nightingale' (*nessuno usignolo)

alcun elmo 'any helmet' (*alcuno elmo)

(b) diciott'/ diciotto uccelli 'eighteen birds'

quest'/ questo amico 'this friend'

quest'/ questa ombra 'this shadow'

nessun'/ nessuna esca 'no bait'

bell'/ bella Italia 'beautiful Italy'

dev'/ deve essere '(s/he) must be'

A possible explanation for this differentiation is given in 4.2.2.

4.2.2 Truncations

4.2.2.1 Specifier Vowel Deletion

In Italian, VD is not always triggered by a vowel clash. For instance, a group of the above-mentioned lexical items obligatorily lose their final vowel in the masculine singular[21] even when the following word begins with a consonant, as is shown in (44).[22] These are precisely the noun specifiers that are listed in (43a), namely *lo* 'the' and its compounds, *uno* 'a, one', *quello* 'that', *bello* 'nice, handsome, beautiful', *alcuno* 'any, some', *nessuno* 'no', *buono* 'good'.[23]

(44) il fiore 'the flower'

un cane 'a dog'

quel bandito 'that highwayman'

bel ragazzo 'handsome boy'

buon pranzo 'good lunch'

nessun gufo 'no owl'

---

[21] The alternation does not take place in the feminine (*la, una, quella,* etc.), probably because the masculine singular is to be considered the morphologically unmarked case, to which the ending *–a* is added to form the feminine. Deletion of the final /o/, contrary to that of the /a/ of the feminine, 'would not suppress unrecoverable information' (Burzio 1987:7).

[22] This process must be distinguished from the above-mentioned elision of section 4.2.1.2.1 because, beside the fact that it only applies to a specific subset of the lexical items that undergo elision, it also yields different output forms (*il* instead of *l'*, *quel* instead of *quell'*, etc.).

[23] Burzio (1987:1-2) actually divides the specifiers into 2 groups: those ending in [n] and those ending in [l] after truncation.

alcun male 'any evil'

4.2.2.1.1 Syntactic and lexical conditioning

The application of this process, which is a type of truncation and which Nespor (1990) calls Specifier Vowel Deletion (SVD), is heavily conditioned by syntactic and lexical factors. It depends on the grammatical function that the relevant word carries out: as can be seen in (45), deletion only takes place when this functions as a specifier (45a, c); it is blocked when such a word has a pronominal function (45b, d).

(45) a. Ho letto *un* lunghissimo libro. (* uno)

'(I) have read a very long book'

b. Ne ho letto *uno* lunghissimo. (*un)

'(I) have read a very long one'

c. Dammi *quel* quadro. (*quello)

'Give me that picture'

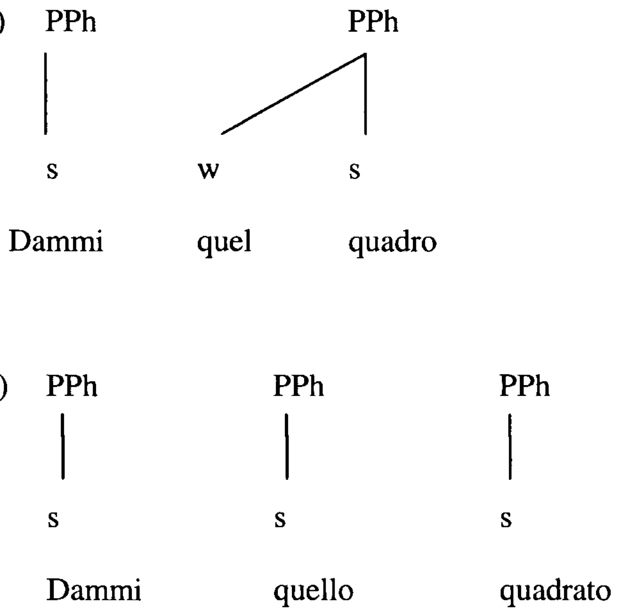
d. Dammi *quello* quadrato. (*quel)

'Give me the square one'

(from Nespor and Scorretti 1985:214-5)

Rizzi (1979) and Vanelli (1979) explain the difference between examples such as those in (45) by referring to phonetically empty syntactic elements. That is, the VD is blocked because of an intervening NP trace (45b) or PRO (45d) immediately following the relevant word. Nespor and Scorretti (1985:215), however, argue that a simpler explanation can be given in terms of prosodic structure: the word that undergoes deletion must be dominated by the weak node of a PPh. *Uno* and *quello*, in (45b) and (45d) respectively, are seen as heads exhaustively forming a PPh (cf. (46)). In fact, Nespor and Scorretti (1985:216-8) show that, on the basis of government theory considerations, the existence of an intervening empty category functioning as head of such PPhs must be ruled out.

(46) (a), (b)



(from Nespor and Scorretti 1985:218)

Because of its limited distribution and of its asymmetric treatment of X-bar categories, Nespor (1990:385, 1993:222-225) concludes that this type of VD must be a case of phrasal allomorphy (see Hayes 1990), by which the different forms are precompiled in the lexicon. She (1993:225) thus defines the frame for the relevant specifiers as in (47).

(47) /[... ____ ... X^o]

4.2.2.1.2 Phonological conditioning

Although a certain lexical conditioning for SVD is undeniable,[24] phonological factors are more important than it may appear at a first glance. A closer look at the items that undergo this process reveals that the choice of these specific words is not arbitrary, but phonologically driven: the noun specifiers listed in (43a) and (44) are all and only those that end in a sonorant consonant when truncated. This restriction is evidently related to a syllable structure constraint of Italian that does not allow non-geminate obstruents in the coda (see 1.3).[25]

As Nespor and Scorretti (1985:215) had already mentioned, we can thus safely conclude that 'all nominal specifiers in Italian obligatorily delete the final /o/ of the masculine singular form', provided the resulting truncated form ends in a sonorant consonant. In a DT framework, SVD could thus be formulated as in (48), where the lexical frame has been included.

(48) *Specifier Vowel Deletion*

$$/o/ \rightarrow [0] \; / \; [\, \ldots \begin{bmatrix} +\text{cons} \\ +\text{son} \end{bmatrix} \_\_\_\_ \; [+\text{cons}] \;]_{[\text{Frame (47)}]}$$

This formulation, however, does not attempt to explain why SVD takes place: it simply describes how it applies. As was mentioned in 3.2, on the other hand, a constraint-based approach may offer a possible explanation. Bullock (1998) suggests that truncation may be triggered by the constraint requiring trochaic feet to be even, namely EVEN-TROCHEE. According to this requirement, uneven (HL) trochees are eliminated whenever possible by deleting their final vowel, and consequently their

[24] See for example a partially fossilized form like *il* as an alternant of *lo* 'the (m.sg.)'.

[25] Even if /s/ appears to form a possible syllable coda in Italian (cf. *pes.ca* 'peach', *mos.tro* 'monster'), it must be banned from word-final position for the following reason. Apart from the above mentioned specifiers, another restricted group of Italian lexemes – among which *santo/ san/ sant'* 'saint', *grande/ gran/ grand'* 'large; great' - undergo a similar truncation process. This time, however, the preceding consonant is involved together with the word-final vowel. According to Burzio (1987:5), nothing would then prevent *questo* from having its final /to/ truncated in exactly the same way. The reason why this does not take place, is precisely because **ques*, exhibiting a final /s/, is not a well formed word, unlike those with a final sonorant.

light syllable.[26] In order for the vowel to be deleted, MAX–V-IO must be outranked by EVEN-TROCHEE. This is in its turn dominated by a constraint allowing only truncated forms ending in sonorants, which we may call *OBS]$_{PW}$ (no obstruents in PW-final position). Such sonorant consonants are preserved because of the constraint MAX–C-IO, by which an input consonant must have a correspondent in the output. Tableaux (49) and (50) contrast the evaluation of the candidates for *quel gatto* 'that cat', in which the specifier undergoes truncation, and *questo gatto* 'this cat', in which truncation cannot take place.

(49)

Input: /kwello gatto/	*[...C_1C_1...]$_\sigma$	SON-DIS	*OBS]$_{PW}$	MAX–C-IO	EVEN-TROCHEE	MAX–V-IO
a. (kwel.lo)(gat.to)					**!	
b. ☞ (kwel)(gat.to)					*	*
c. (kwe:)(gat.to)				*!	*	*
d. (kwell)(gat.to)	*!				*	*
e. (kwel)(lgat.to)		*!			*	*

Notice that, in order to discard candidates d. and e., the constraints *[...C_1C_1...]$_\sigma$, disallowing tautosyllabic geminates, and SON-DIS, requiring a certain sonority distance between the segments forming an onset, must also be introduced. These two constraints must be ranked quite high in Italian, although their ranking with respect to *OBS]$_{PW}$ cannot be established.

[26] Notice that these considerations apply to all items in (44) except for *il* ~ *lo* 'the (m.sg.)' which, being a proclitic, is often unfooted. They do apply, however, when a preposition is attached to such an article,

(50)

Input: /kwesto gatto/	$*[...C_1C_1...]_\sigma$	SON-DIS	$*OBS]_{PW}$	MAX–C-IO	EVEN-TROCHEE	MAX–V-IO
a. ☞ (kwes.to)(gat.to)					**	
b. (kwes)(gat.to)			*!	*	*	*
c. (kwe:)(gat.to)				*!*	*	*
d. (kwest)(gat.to)			*!		*	*
e. (kwes)(tgat.to)		*!	*		*	*

The tableau in (50) shows that MAX–C-IO must outrank EVEN-TROCHEE, in order for candidate a. to be more harmonic than candidate c.

Another characteristic of SVD is that it is blocked when the following word begins with an [s] followed by a consonant (so called '*esse impura*') (see 2.2) or with one of the consonants that – in SI – are always long (i.e. [ɲ], [ʎ], [ʃ], [ts], [dz], [j][27]). What such consonants and /sC/ have in common is that they cannot be exhaustively contained in a syllable onset, since in Italian geminates must be heterosyllabic and onset /sC/ clusters would violate Sonority Sequencing. Therefore, the final vowel of the preceding specifier is maintained in order for the 'esse impura' or the first segment of the geminates to be resyllabified as codas (Burzio 1987:4; Nespor 1993:177-178), as in (51).[28]

(51) a. nessuno studente 'no student' → [nes.su:.nos.tu.dɛn.te]

b. bello zaino 'nice rucksack' → [bɛl.lo**dz**.dzai̯.no]

In an OT perspective, these forms violate the alignment constraint ALIGN-L (PW,σ), since the left edge of the PWs [studɛnte] 'student' and [ddzaino] 'rucksack' does not

as in *nel* ~ *nello* 'in the (m.sg.)' or *sul* ~ *sullo* 'on the (m.sg.)'.

[27] Onset [j], unlike the other palatals, is only doubled word-initially.

[28] It must be noted, however, that /sC/, /ts/, /dz/, /ɲ/, /ʎ/ and /ʃ/ can actually be found in absolute initial position in Italian, since forms like *scuola* [skwɔla] 'school' or *zio* [tsio] are well formed in isolation. In such cases, initial /s/ must be syllabified together with the following consonant, while the geminate consonants, whose first segment does not find a coda slot to be syllabified in, must be degeminated. Burzio (1987:4), however, suggests that tautosyllabicity of /sC/ and degemination apply only as a last resort.

coincide with the left edge of a syllable. ALIGN-L (PW,σ) must therefore be outranked by other constraints, such as *[...C_1C_1...]$_\sigma$ or SONSEQ (see tableaux (52) and (53) below). On the other hand, neither of the optimal forms exhibits more violations of EVEN-TROCHEE than the suboptimal truncated candidates. This latter constraint will therefore be left out of tableaux (52) and (53), reporting the evaluation of the candidates for the phrases in (51).

(52)

Input: /nessuno studɛnte/	*[...C_1C_1 ...]$_\sigma$	SON SEQ	*COMPLEX ^{COD}	MAX -μ- IO	MAX -C- IO	ALIGN-L (PW,σ)
a. [nes.sun.stu.dɛn.te]		*!				
b. ☞ [nes.su:.nos.tu.dɛn.te]						*
c. [nes.suns.tu.dɛn.te]			*!			
d. [nes.sun.<s>tu.dɛn.te]					*!	

(53)

Input: /bɛllo ddzai̯no/	*[...C_1C_1...]$_\sigma$	SON SEQ	*COMPLEX^{COD}	MAX -μ- IO	MAX -C- IO	ALIGN-L (PW,σ)
a. [bɛl.ddzai̯.no]	*!					
b. ☞ [bɛl.lodz.dzai̯.no]						*
c. [bɛldz.dzai̯.no]			*!			
d. [bɛl.dzai̯.no]				*!		

In tableau (52), the truncated candidates are ruled out by ranking the constraints SONSEQ (violated by candidate a.), *COMPLEX^{COD} (violated by candidate c.) and MAX-C-IO (violated by candidate d.) above ALIGN-L (PW,σ) (violated by the optimal candidate b.). In tableau (53), the suboptimal candidate a., in which a geminate consonant is exhaustively syllabified in an onset, violates *[...C_1C_1...]$_\sigma$, while candidate d. violates MAX-μ-IO by degeminating the double consonant, thus losing a mora. Therefore, these two constraints are also ranked above ALIGN-L (PW,σ), violated by the optimal candidate. Ranking among the dominant constraints cannot be established on the basis of these data, although $*[...C_1C_1...]_\sigma$ is probably ranked above all others, since it is the only one that seems to be inviolate in Italian.[29]

4.2.2.1.3 The plurals

The issue is further complicated by the fact that, for the lexical items having /l/ as the final consonant of the truncated form, the masculine plural patterns in a peculiar way, too. In particular, in isolation such items display forms of the type *quelli* [kwelli]/ *belli* [bɛlli], before /sC/, /dz/, /ts/, /ʃ/, /ɲ/, /ʎ/ and all vocoids forms of the type *gli* [ʎi]/ *quegli* [kweʎʎi]/ *begli* [bɛʎʎi] and before other consonants forms of the type *i* [i]/ *quei* [kwei]/ *bei* [bɛi] (cf. (54)).

(54) (a) Questi ragazzi sono belli

'These boys are good looking'

(b) Bei ragazzi

'Good looking boys'

(c) Begli studenti

'Good looking students (m.)'

(d) Begli amici

'Good looking friends (m.)'

(from Burzio 1987:5)

Examples such as those in (54) evidently constitute a problem for the above accounts. For example, both the forms *belli* and *begli* end in a vowel, and could thus incorporate the first segment of a geminate or an /s/ in the coda of their final syllable.

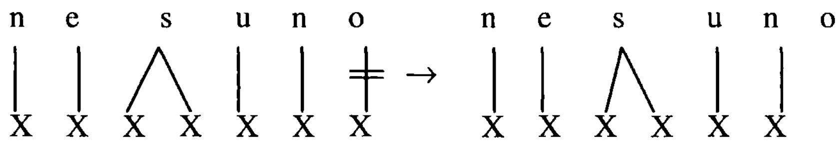
---

[29] SONSEQ and MAX-μ-IO are violated in isolated forms such as *studente* [stu.dɛn.te] 'student' or *zaino* [dzai̯.no] 'rucksack', *COMPLEX^{COD} in loanwords such as *trust* 'trust', MAX-C-IO in fast speech forms in which consonant clusters are simplified.

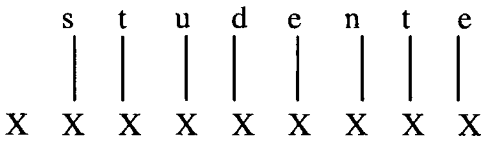
4.2.2.1.4 Burzio's proposal

According to Burzio (1987), the notion that contexts like /sC/ simply block truncation is inadequate. The solution he offers is based on the idea that there are actually two processes applying to these forms: a process of 'prosodic reduction' affecting the noun specifiers, and one of 'augmentation' affecting words beginning with /sC/ or one of the geminate consonants. The first operation deletes the final timing slot of words like *nessuno* or *bello*, thus corresponding to truncation; the second inserts a timing slot (and probably a whole syllable) at the beginning of words like *studente* or *zaino*, in order to rescue the unsyllabifiable floating segment. When a word with 'prosodic reduction' and one with 'augmentation' are put together, the floating vowel goes to fill in the empty timing slot; this is called by Burzio (1987) 'pseudo-epenthesis'. This interaction is shown in (55), exemplified on *nessuno studente* 'no student'.

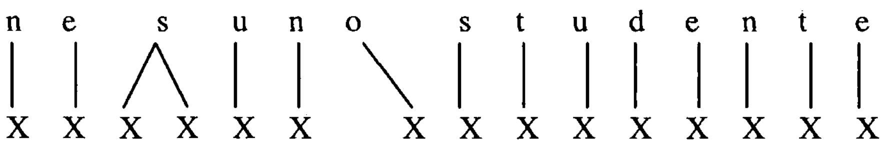
(55) (a) PROSODIC REDUCTION



(b) AUGMENTATION



(c) PSEUDO-EPENTHESIS



(from Burzio 1987:6)

The final /o/ of words like *nessuno* cannot be syllabified, and is thus suppressed by general convention, unless the following word exhibits a free time unit. Such a unit, in its turn, will only be filled – and obligatorily so - when a floating vowel is available (cf. 5.2).

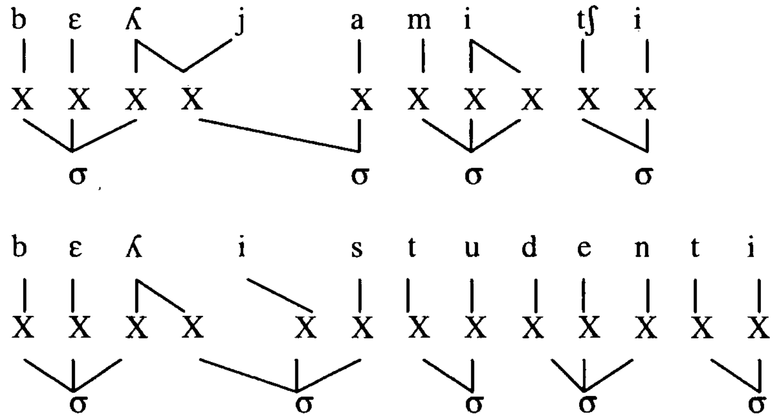
Burzio (1987:10) further argues that obligatoriness of prevocalic elision is also determined by prosodic reduction. In fact, as was shown in (43a), elision is obligatory precisely for the same group of forms which undergo truncation (i.e. the masculine singular sonorant-final specifiers). The forms in (43b), on the other hand, are simply affected by the optional connected speech process of VE exemplified in (38), although they are more fossilized than others.

> PR [Prosodic Reduction] will thus occur regardless of phonological environment, so long as the relevant syntactic conditions are met. The phonological environment will then interact with syllabification principles in determining the various different outcomes, such as 'elision', 'truncation' and 'pseudo-epenthesis'
>
> (ibid.:11)

As regards the masculine plural, Burzio (ibid.:7-10) proposes that, as a result of prosodic reduction, the final /I/ turns into a [j] and that, consequently, the preceding [l] is palatalized to [ʎ]. In fact /I/, unlike the /o/ of the singular, does not need to be left floating, because it can lose its syllabicity and become /j/ when prosodic reduction applies. However, by itself *begli* is not well formed, because it lacks a final vowel. Therefore, it can only appear in two contexts: before a vowel, which can function as a peak for the ill-formed syllable (56a), and in the context of pseudo-epenthesis, where the final vocoid can reacquire its syllabicity by associating with the empty time unit (56b).

(56) (a)

(b)



(from Burzio 1987:8)

Burzio's (1987) account of the processes yielding forms like *bei, quei, i*, however, is less convincing. It is based on the notion that, while palatalization of a geminate [ll] yields [ʎʎ], palatalization of a simple [l] yields [i]. In the case of *bei* (rather than *begli*), palatalization would apply before rather than after gemination.[30]

Even in its most successful intuitions, Burzio's (1987) explanation has more validity at the diachronic than at the synchronic level. Burzio himself (fn.8, p.3) admits that, in varieties of Italian where [ts], [dz] and the palatal consonants are not geminated, the behaviour of the specifiers with respect to truncation must in part be learned as an idiosyncrasy. At least for such varieties, therefore, his account must be integrated with Hayes (1990) and Nespor's (1990) notion of phrasal allomorphy precompiled in the lexicon.[31]

4.2.2.2 Verb Final Vowel Deletion

Another truncation process of (especially Northern) Italian affects only words belonging to a specific syntactic category: that of verb. It has therefore been labelled as Verb Final Vowel Deletion (henceforth VFVD) (NV:32). Such an operation optionally deletes a word-final /e/ or /o/ if such a vowel is unstressed, and immediately preceded by a coronal sonorant consonant, preceded, in turn, by a vowel (Nespor 1990:385, 1993:83). Moreover, as was mentioned above, the word affected by the deletion must be a verb, and the following word must be included in the same specific domain. The process is illustrated in (57), where it is apparent that the deletion can affect verbs (57a) but not other constituents (in this case, nouns) (57b).

---

[30] This account is all the less convincing considering that geminate consonants are generally seen as underlying in Italian (cf. 3.2).

[31] Given the (often necessarily) stipulatory nature of this explanation, it is particularly difficult to express it in OT terms. An OT account of the plural patterns will thus be left out.

(57) (a) So che vuol[0] nuotare. (< vuole '(s/he) wants)

'(I) know that (s/he) wants to swim'

(b)*Ho le suol[0] nuove. (< suole 'soles')

'(I) have new soles'

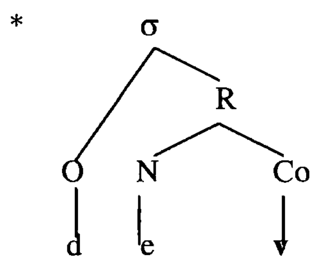
(c)*So che dev[0] nuotare. (< deve '(s/he) must')

'(I) know that (s/he) must swim'

(from NV:32-3)

(57c) also shows that VFVD cannot take place if the relevant vowel is not preceded by a sonorant consonant. This is clearly due to the above-mentioned modified Coda Filter, by which in Italian a syllable coda can only contain a sonorant consonant, the first segment of a geminate or, when word-internal, an /s/. It is thus not possible to parse a sequence like [dev] into one self-contained syllable, since the process would yield a disallowed representation, as in (58).

(58)



The deletion seems to apply regardless of whether the following word begins with a vowel or with a consonant; therefore, it has nothing to do with vowel-clash resolution.

(59) (a) Parlan piano. (< parlano '(they) speak')

'(They) speak low'

(b) ?Parlan ancora.

'(They) are still speaking'

However, forms like that in (59a), though not obligatory, are frequent at all speech rates, while forms like that in (59b) are less frequent and only seem to appear in fast speech (Nespor 1993:227). If we consider a case like *Parleran piano* '(They) will speak low' vs. *Parlerann'ancora* '(They) will speak again', moreover, it looks as if the optional deletion process that applies before a vowel is not the same as the one applying before a consonant: it is the above-mentioned fast speech clash-avoiding VD. The difference can be detected in the doubling of the [n] and perhaps in the fact that the second segment of such a geminate is resyllabified into the following syllable. By the same token, *Parlan'ancora* will be different from *Parlan piano*. In the first case, we can talk about 'vowel elision', in the second case about 'truncation' (see also Nespor 1993:227).

Nespor (1990:393, 1993:228) argues that, although it applies only to certain forms, VFVD cannot be described in terms of syntactic constituents. It only makes reference to the lexical category of verb, but such a verb does not have to be part of a VP; it may also, for instance, have a nominal function, as in (60).

(60) [Questo continuo mangiar[0] con le mani]$_{NP}$ mi disturba. (< mangiare)

'This continuous eating with our hands bothers me'

(from Nespor 1993:228)

On the contrary, it takes place within a specific prosodic domain, which can be identified as I, on the basis of examples such as those in (61).

(61) (a) [Se non vuol[0/e] partire,]$_I$ [non partirà.]$_I$

'If (s/he) does not want to leave, (s/he) will not leave'

(b) [Se non vuol[e],]$_I$ [non partirà.]$_I$ (*vuol[0])

'If (s/he) does not want to, (s/he) will not leave'

Nespor (1990:386) points out that this feature, together with others such as sensitivity to speech rate and to pauses and insensitivity to the presence of phonetically empty syntactic constituents, would seem to characterize VFVD as a prosodic operation.

However, its asymmetric treatment of X-bar categories is instead typical of Hayes's (1990) lexically precompiled phonological processes.

Nespor (ibid.:392) thus concludes that VFVD is a rule of phrasal allomorphy which can be represented as in (62).

(62) $$\begin{bmatrix}-\text{cons}\\ -\text{high}\\ -\text{back}\end{bmatrix} \rightarrow \quad 0\ /\ [\ldots \begin{bmatrix}+\text{cons}\\ +\text{son}\\ +\text{cor}\end{bmatrix} \_\_\_\_ \ [+\text{cons}]\ ]_{\text{FRAME 1}}$$

Frame 1:

$[\ \ldots\ [\ \ldots \_\_\_\_\ ]_{\text{VERB}} \ldots X\ ]_{\text{I}}$ $\qquad X = 0$

Sensitivity to speech rate and pauses and insensitivity to phonetically empty elements would thus be a consequence of the fact that VFVD has a prosodic, rather than a syntactic, frame.[32]

Unlike SVD, VFVD applies before any type of consonants and consonant clusters, including /sC/, /ts/, /dz/ and the palatals, as the examples in (63) show.

(63) (a) *son sporchi* [sonspɔrki] '(they) are dirty' (< *sono*)

(b) *mangiar gnocchi* [mandʒarɲɔkki] 'to eat gnocchi' (< *mangiare*)

According to Nespor (1993:230), the contrast between the two processes is due to the fact that VFVD, having a prosodic frame, is not necessarily subject to phonotactic constraints, by which SVD, with its syntactic frame, is instead conditioned.[33]

[32] It must be pointed out, however, that the process only applies to certain verb forms, such as auxiliary and modal verbs, the 3rd person plural ending *–no*, the infinitive endings *–are/ -ere/ -ire*. There are other forms which, although ending in vowel + coronal sonorant + mid vowel, cannot undergo truncation (e.g. *lavoro/ *lavor* '(I) work'). Moreover, in some Italian varieties the first person plural ending *–mo* can also be truncated, although its penultimate segment is not coronal.

[33] In Burzio's (1987) analysis, instead, such a distinction is accounted for by assuming that in the case of the specifiers prosodic reduction is not accompanied by deletion of the vocoid at the segmental level, while in the case of the verbs in (63) the deletion takes place both at the skeletal and at the segmental

This differentiation is instead problematic for an OT analysis of the phenomena, since the segmental sequences in phrases such as *son sporchi* '(they) are dirty' and *uno (*un) sporco* 'a dirt' are virtually identical. Although the truncated forms *son* and *un* may both be triggered by the constraint EVEN-TROCHEE, at least one of the anti-truncation constraints SONSEQ and *COMPLEX^{COD} should be dominated by ALIGN-L (PW,σ) in the first case,[34] while they should both outrank it in the second. That is, if the ranking postulated in (52) in order to account for forms such as *nessuno studente* and *uno sporco* is maintained, a candidate like *son sporchi* will never be selected. As is shown in tableau (64), the winning candidate will always be *sono sporchi*, whatever the syllabic parsing of *son sporchi*.

(64)

Input: /sono spɔrki/	SONSEQ	*COMPLEX^{COD}	ALIGN-L (PW,σ)
a. [son.spɔr.ki]	*!		
b. ☞ [so:.nos.pɔr.ki]			*
c. [sons.pɔr.ki]		*!	*

More generally, it is problematic, within a constraint-based approach, to illustrate phenomena that behave asymmetrically with respect to different lexical categories.

4.2.2.3 Truncated nouns

At least for certain forms, truncation is a completely productive process as far as verbs are concerned. This is not the case with the other lexical categories, where nevertheless some sporadic fossilized instances of truncation are found. Such forms comprise a restricted group of nouns referring to male human beings such as *signore* 'mister/ gentleman', *dottore* 'doctor (m.)', *generale* 'general' or *professore* 'professor (m.)', plus another few nouns such as *mare* 'sea' or *colore* 'colour' (Nespor 1990:387-8, 1993:84). Although they are exempt from phonological conditioning, such nouns

level. Consequently, no segment is available to fill in the empty position of the pseudo-epenthetic forms.

[34] This should actually be a case of free ranking, since the form *sono sporchi* is also acceptable.

only undergo truncation in certain syntactic contexts, in which they are generally used as titles, as is shown in (65a) as opposed to (65b).

(65) (a) *signor Bianchi* 'mister Bianchi', *dottor Brambilla* 'doctor Brambilla', *mar Ligure* 'Ligurian sea', *color petrolio* 'petroleum colour'

(b) **un signor bianco di capelli* 'a white-haired gentleman', **un dottor bravissimo* 'a very good doctor', **un mar limpidissimo* 'a very clear sea', **un color per dipingere le scale* 'a colour to paint the stairs'

(from Nespor 1993:84)

Just like the above-mentioned verbs, these items do not alternate according to the initial sound(s) of the following word (e.g. *Signor Carini/ Signor Aloisio/ Signor Stoppani* vs. *quel carro* 'that cart'/ *quell'alloro* 'that laurel'/ *quello stormo* 'that flock'). Contrary to VFVD, however, this process is not productive and is limited to a restricted number of lexemes; therefore, these forms must be precompiled in the lexicon, being the result of diachronic rather than synchronic truncation.

4.2.3 Syncope

Contrary to elision and truncation, syncope is not common in Italian, for at least two related reasons. Firstly, the constraint against monomoraic syllables, although absolute in the case of stressed syllables (cf. 3.2), appears not to be operative – i.e. to be ranked low under an OT perspective – for unstressed ones. Secondly, the range of permitted consonant clusters is very limited and consonants cannot be syllable peaks. Syncope only applies in fast speech, as it does in many other languages. As was specified in section 1.2, fast speech phenomena are not within the scope of this thesis.

4.3 Syncope and other deletion processes in Arabic

Contrary to Italian, in Arabic syncope is by far the most common – and most studied – deletion process affecting vocoids. Therefore, the largest part of this section (i.e. subsection 4.3.2) will be devoted to the analysis of this phenomenon, of its interaction with factors such as vowel quality and stress placement, and of the complex cross-dialectal differences in its implementation. Before tackling this all-important issue, however, a brief account of Glide Deletion will also be given (subsection 4.3.1). The last subsection will instead show that vowel elision and truncation are not common in AA.

4.3.1 Glide Deletion

In Arabic, as was anticipated in 2.3, deletion does not only affect the vocoids on the affixal tier, but also those on the root tier. In particular, a root vocoid is often deleted in intervocalic position, as shown in (66), where all examples are from AA.[35] This is traditionally called Glide Deletion (GD); I will retain this terminology although, as was said in 2.3, I consider all Arabic high vocoids to be underspecified for syllabicity underlyingly.

(66) /ramaIa/ → [rama] '(he) threw'

/sammaIa/ → [samma] '(he) named'

/naIam/ → [na:m] '(he) slept'

/ʃaIal/ → [ʃa:l] '(he) carried'

/ʃaUaf/ → [ʃa:f] '(he) saw'

/naIIl/ → [na:l] '(he) obtained'

/TaUUl/ → [Ta:l] '(he) grew taller'

/maUat/ → [ma:t] '(he) died'

/baIaDat/ → [ba:Dat] '(she) laid eggs'

[35] GD is followed by Vowel Contraction (yielding a long vowel) and, in dialects such as AA, by Final Vowel Shortening (see 3.3).

/baUab/ → [ba:b] 'door'

/naIab/ → [na:b] 'canine tooth'

/naUadi/ → [na:di] 'club'

Examples such as *mayaadiin* [maja:di:n] 'squares', *Tawiil* [Tawi:l] 'tall' or *samaawi* [sama:wi] 'azure', which do not undergo GD, appear to indicate that a root vocoid is only deleted when the surrounding vowels are short.

In a DT perspective, therefore, the process may be represented as in (67).

(67)

$$\begin{array}{lccc} & X & & X \\ & | & & | \\ G \rightarrow 0 \; / & V & \_\_\_\_ & V \end{array}$$

where G = root vocoid, V = affixal vocoid

GD is easily motivated on the basis of sonority distancing considerations: the sonority of the root high vocoid being very close (or identical) to that of the surrounding vowels, the trough between them is insufficient and thus triggers deletion of the glide and contraction of the vowels into one syllable (Angoujard 1990:54).[36]

In an OT approach, these considerations may be translated into the trigger-constraint $*\breve{v}G\breve{v}$, militating against the presence of glides between short vowels. Clearly, this should outrank MAX-IO, the anti-deletion constraint.

GD, however, does not apply to all lexical items in which the triggering environment is met. Words such as *shawarma* 'doner kebab', *sawa* 'together', *dawa* 'medicine', *duwali* 'international', for instance, preserve their glide in spite of its being collocated between two short vowels. Only a restricted set of morphological classes regularly undergoes GD. In particular, this process characterizes trilateral verbs which have a high vocoid as a second or third radical. These are traditionally called 'hollow verbs' (e.g. *shaaf* 'see', *naam* 'sleep') and 'final-weak verbs' (e.g. *rama* 'throw', *nisi* 'forget') respectively.

Due to the lexical and morphological restrictions on its distribution, some scholars (cf. for instance Watson 2002) consider GD a diachronic phenomenon. However, the existence of alternants displaying a high vocoid (see (68)) seems to indicate that such high vocoids are still present underlyingly.

(68) baab 'door' ~ 'abwaab 'doors'

naab 'canine tooth' ~ 'anyaab 'canine teeth'

shaaf '(he) saw' ~ shuft '(I/you m.sg.) saw'

shaal '(he) carried' ~ shilt '(I/you m.sg.) carried'

rama '(he) threw' ~ rameet '(I/you m.sg.) threw'[37]

Because of its limited distribution, though, VD must be an allomorphy operation restricted to verbs and applying within the stem.[38] The allomorphs of nouns such as *baab* 'door' or *naab* 'canine tooth' will instead be directly precompiled in the lexicon.

4.3.2 Vowel syncope

As was mentioned above, syncope of short vowels in open syllables is by far the most common vocoid deletion process in Arabic. It takes place in all dialects,[39] though to different extents and under different conditions.

[36] A similar phenomenon takes place in English words such as *fire* [fajə] ~ [fa:] or *hour* [awə] ~ [a:].
[37] In dialects such as AA, the [e:] of the first and second persons is derived from /ai̯/ by an independently motivated process of monophthongization (see 2.3).

[38] Harrell (1957:67) points out that GD must be an early process, applying in the lexicon: when a postlexical process applies, if two identical vowels come into contact they do not seem to contract, but to remain in separate syllables. This is the case of h-deletion in Egyptian Arabic, as shown in (ii) below.

(ii)	9andaha	'she has'	→	9an.da.a
	laha	'to her'	→	la.a
	ismuhum	'their name'	→	is.mu.um

[39] A more limited process of syncope also applies in Classical Arabic, where word-medial short vowels are deleted only when between two consonants that are easily assimilated, as in *masasa* > *massa* 'he touched', *raadid* > *raadd* 'returning' (Cantineau 1960:108). This change, however, appears to have diachronic rather than synchronic significance.

Examples from AA are given in (69), where alternation with non-syncopated forms is also shown.

(69) *shirib* '(he) drank' ~ *shirbu* '(they) drank'

fihim '(he) understood' ~ *fihmat* '(she) understood'

kibir '(he) grew up' ~ *kibru* '(they) grew up'

SaaHib 'friend (m.)' ~ *SaaHba* 'friend (f.)'

btaaxud '(you m.sg.) take' ~ *btaaxdi* '(you f.sg.) take'

faahim 'understanding (m.sg.)' ~ *faahma* 'understanding (f.sg.)'

taakul '(you m.sg.) eat' ~ *taaklu* '(you pl.) eat'

saafir 'travel (m.sg.)!' ~ *saafri* 'travel (f.sg.)!'

bitraasil '(you m.sg.) correspond' ~ *bitraaslu* '(they) correspond'

Other forms, though lacking an alternant to show complementary distribution of V ~ 0, often exhibit free variation between the two (70).[40]

(70) *muHammad* ~ *mHammad* 'Mohammed'

duru:s ~ *dru:s* 'lessons'

Himaar ~ *Hmaar* 'donkey'

bujuut ~ *bjuut* 'houses'

kilaab ~ *klaab* 'dogs'

kitaab ~ *ktaab* 'book'

In all the above examples, the deleted vowels are in monomoraic open syllables. In fact, short vowels in closed syllables do not generally undergo syncope. This trend is particularly clear in the examples in (69), where the target vowel is in a closed syllable in the non-syncopated alternants, while in the syncopated alternants such a syllable has been opened by the addition of a vowel-initial suffix after its coda consonant.

[40] However, there is a restricted number of commonly used lexical items, such as *ktiir* 'a lot' or *kbiir* 'big', in which the vowel never seems to surface at a normal speech rate; it is likely that, in these cases,

These facts can be accounted for by means of Broselow's (1992) Bimoraicity Constraint, requiring that syllables be maximally and optimally bimoraic: in a language like Arabic, suboptimal monomoraic syllables are eliminated whenever possible. As was mentioned in 1.5, in OT the constraint $*\sigma_{\mu}$ militates against monomoraic syllables, triggering their deletion.

In the North African dialects, however, a vowel may undergo syncope even if it is in a closed syllable, provided the preceding and following consonants form a pronounceable cluster, as in *Hanak* > *Hank* 'jaw, cheek' or *qalam* > *qalm* 'pen' (Cantineau 1960:107).[41]

Other cross-dialectal differences concern the quality of the target vowel, the influence of stress placement and the triggering environment.

4.3.2.1 Quality of the target vowel

It is not by chance that, in all the examples in (69) and (70), the deleted vocoid is either /I/ or /U/. As was specified in 4.1, high vowels, being less sonorous, are generally more prone to undergo syncope than low vowels are. Cantineau (1960:108) points out that a short [a] is actually longer than a short [i] or [u], which is actually still related to a difference in sonority.

In all Arabic varieties, in fact, closed or semiclosed vowels such as [i], [u] or [ə] are more readily deleted than open vowels such as [a] and [ɛ]. However, while some dialects completely ban syncope of non-high vowels, others allow it, and to different degrees.

In certain varieties, syncope affects all short vowels, regardless of their degree of sonority. In the Maghreb , for instance, all three short vowels undergo syncope in open syllables with the same frequency,[42] as is shown by the Tunisian examples in (71).

---

the deleted forms have been lexicalized. Kenstowicz (1981:25) actually appears to assume that, in Palestinian Arabic, even words such as *ktaab* 'book' have undergone diachronic syncope.

[41] According to Cantineau (1960:119), in these dialects all short syllables have disappeared, that is to say that all short vowels in open syllables have been deleted.

(71) *fras* 'mare' < *faras-*

qluub 'hearts' < *quluub-*

klaab 'dogs' < *kilaab-*

(Cantineau 1960:110)

/a/ is also deleted alongside high vowels in Muscat (Zawaydeh 1997), San9ani (Watson 1999a, 2000, 2002) and Iraqi Arabic (Majdi and Winston 1994:197).

In varieties such as Cairene, on the other hand, syncope only applies to high vowels. In the Levant, most dialects belong to this second category, i.e. they tend to preserve /a/ as opposed to /I/ and /U/, which undergo syncope. In most of the urban dialects of this area (Cantineau 1960:109), including AA, the majority of short unstressed low vowels in open syllables are preserved, both in pretonic and posttonic position (see (72)).

(72)	*maHátta*	'station, bus-stop'
	karíim	'noble, generous'
	madíina	'city'
	máktaba	'library'
	bádanu	'his body'
	báladi	'my village'
	bá'ara	'cow'

[42] In Libyan Arabic, however, low-vowel syncope is subject to restrictions. According to Abumdas (1985:194), in this dialect the environment for syncope should contain a high vowel in an open syllable, but it is not necessarily the high vowel that is deleted: it may be the preceding low vowel.

In these same dialects, however, when a word contains three consecutive open syllables, all three with the vowel /a/, one of these is syncopated (the examples in (73) are, as usual, from AA).[43]

(73) /raqabat + u/	→	[raʔbatu]	'his neck'
/xaʃabataIn/	→	[xaʃabteen]	'two pieces of wood'

However, in AA there are cases of (very common) lexical items in which an /a/ is deleted even when it is not preceded by other /a/s, as in /ʒadi:d/ > [ʒdi:d] 'new'.[44]

Kenstowicz (1981:25-6) also reports that many Levantine dialects have morphologically conditioned /a/-syncope. According to him, the most restricted environment is found in Damascene Arabic, where it only applies to 'the second syllable of C(C)aCaC verb stems before the 3 sg. f. perfect suffix –at' (p. 25). Examples that also fit AA are in (74).

(74) katab + at	→	katbat '(she) wrote'
nsaHab + at	→	nsaHbat '(she) withdrew'
shtaghal + at	→	shtaghlat '(she) worked'

(from Kenstowicz 1981:26)

In AA, as in the dialect of Horan, the process is extended to apply before the feminine noun suffix –at in construct, as in the examples in (75).

(75) samaka 'one fish' → samkati 'my fish', samkatu 'his fish'

[43] According to Cantineau (ibid.), in Damascene Arabic it is generally the third /a/ that is deleted, while the second one often undergoes reduction.

ba'ara 'cow' → ba'rati 'my cow' (but *ba'araat* 'cows')

In AA, as in Cairene, an /a/ may also be deleted within the PW after a CVVC sequence and before a vowel-initial suffix (Watson 2000:19), as in /baniʔa:dam-i:n/ > [beniʔadmi:n] 'human beings'.

For dialects in which syncope only applies to high vowels, this will have to be specified in the structural description in a DT representation. Within OT, on the other hand, restrictions on the quality of the deleted segment will be expressed by means of constraints banning specific short vowels from open syllables and of their ranking with respect to MAX-V-IO. In cases like AA, the constraints NO-i]$_\sigma$ and NO-u]$_\sigma$, disallowing high vowels in light syllables, will be ranked above MAX-V-IO, while NO-a]$_\sigma$ – stating that /a/ is not allowed in light syllables – will be ranked below it.[45]

4.3.2.2 Stress assignment

As was mentioned above and in 4.1, the distribution of stress is cross-linguistically a key factor in determining the implementation of vowel syncope: unstressed vowels, being weaker, are more prone to undergo deletion than stressed vowels. In fact, in most Arabic varieties only unstressed vowels are syncopated. AA, like Cairene and other urban dialects, is one of these, as the examples in (76) illustrate.

(76) *fíhim* ['fihim] '(he) understood' (*[fhim])

rúkab ['rukab] 'knees' (*[rkab])

Dúhur ['Duhur] 'noon' (*[Dhur])

mítil ['mitil] 'like' (*[mtil])

---

[44] Kenstowicz (1981:28) suggests that /a/ may have gone through the intermediate stage of a reduction to |i| before deleting.

[45] As was said in 4.1, the ranking IDENT-IO[a] » $*\sigma_\mu$ » IDENT-IO[i] obtains the same result and is perhaps more transparent. However, constraints such as NO-i]$_\sigma$ will be used as shorthand.

In the rural dialects of Jordan, Oran and part of Palestine, on the other hand, even high vowels which are originally stressed seem to undergo syncope, as in *rkab* (< *rúkab*) 'knees', *lHa* (< *líHa*) 'beards' (Cantineau 1960:109). San9ani, too, exhibits deletion of underlyingly stressed vowels, particularly in pre-pausal position (Watson 2000:17). This, however, may be a consequence of the fact that, in connected speech, stress in San9ani Arabic is often subject to fluctuations (Watson 2002:73).

Even in dialects in which stressed vowels are never deleted, however, syncope interacts with stress assignment in interesting ways.[46] Conspicuous literature has been devoted to the reciprocal relation between these two phenomena, especially with respect to a set of data from Palestinian - and generally Levantine – Arabic (see for example Brame 1974, Kenstowicz 1981, Kiparsky 2000). In these dialects, short high vowels are not always syncopated in open syllables, even when unstressed. As the examples in (77) show, a contrast is actually exhibited by verbs with high stem vowels between forms with subject suffixes and forms with object suffixes.

(77) a. /fihim/ → ['fihim] 'he understood'

/simiʕ/ → ['simiʕ] 'he heard'

b. /fihim + na/ → ['fhimna] 'we understood'

/simiʕ + na/ → ['smiʕna] 'we heard'

c. /fihim + na/ → [fi'himna] 'he understood us'

/simiʕ + na/ → [si'miʕna] 'he heard us'

(from Brame 1974:43)

While in the base form (77a) the first /i/ is always preserved due to the fact that it is stressed, the other two forms, though apparently identical, exhibit opposite behaviour with respect to syncope: while in the form with the subject suffix (77b) the /i/ is deleted, in that with the object suffix (77c) it is preserved. Since the stress algorithm

[46] Further interaction with epenthesis will be dealt with in chapter 5.

(cf. 1.4) should not assign prominence to the target vowel in any of the latter two forms, such a vowel should undergo syncope in both cases.[47]

Within a DT framework, this apparent puzzle can be traced back to the fact that, as was mentioned in 1.2.5 and 3.3, subject suffixes belong to an inner stratum with respect to object suffixes. Therefore, *fhimna* 'we understood' would derive from [fihim-na], while *fihimna* 'he understood us' would derive from [[fihim]na]. Brame (1974) further assumes that stress assignment applies cyclically, i.e. first to the inner constituent [verb+subj.] and then to the more inclusive [[verb+subj.]obj.] (also see Kenstowicz 1981:23). Therefore, the first syllable in [[fihim]na] does not syncopate because it bears the prominence assigned to it in the first cycle (see (78)).

(78) *Interaction of cyclic stress assignment and syncope* (from Brame 1974:44)

	a. [fihim]	b. [fihim + na]	c. [[fihim]na]
	'he understood'	'we understood'	'he understood us'
1st cycle			
Stress ass.	fíhim	fihímna	fíhim
2nd cycle			
Stress ass.	-	-	fihímna
Syncope	-	fhímna	-
	fíhim	fhímna	fihímna

The table in (78) shows that, when stress assignment applies for the first time, subject agreement endings have already been attached, but object suffixes have not. Therefore, the form with the object suffix (c), contrary to that with the subject ending (b), receives stress on the first vowel. In the second cycle, the object suffix has been

[47] An analysis by which syncope is blocked in cases like *fihimna* 'he understood us' in order to maintain the functional distinction with *fhimna* 'we understood' is refuted by Kenstowicz (1981:22) on the basis of two observations. Firstly, the process is blocked even when it would not lead to confusion of two separate forms (e.g. *fihimkum* 'he understood you (pl.)'). Secondly, syncope sometimes does lead to homophony between different forms (e.g. *fihmu* 'they understood; he understood him; his understanding').

added to the form in (c), and thus prominence is assigned to the second vowel. Consequently, the stress assigned in the first cycle is reduced, but still prevents the application of syncope.

Kenstowicz (1981:24-5) notes that syncope is also blocked in noun stems when possessive suffixes are added, which confirms that these, too, must be assigned to level 3. The examples in (79) are from AA.

(79) *birákna* 'our pools' (**brakna*)

rukábna 'our knees' (**rkabna*)

ghiráana 'our glue' (**ghraana*)

In an OT framework, the postulation of intermediate levels of representation is ruled out. An account of the facts illustrated above, however, can be based on correspondence between surface forms, i.e. OO-correspondence (Kager 1999:281-7). The OO-constraint HEADMAX-BA, by which every segment in the base's prosodic head - i.e. every segment that is stressed in the base - has a correspondent in the affixed form will thus account for the preservation of the unstressed vowel in the accusative forms. For instance, since in the base form *fíhim* 'he understood' the first vowel is stressed, its correspondent in the affixed form *fihímna* 'he understood us' cannot be deleted. In order to obtain this result, HEADMAX-BA will have to dominate NO-i]$_\sigma$, which in its turn, as was stated in 4.3.2.1, outranks MAX-V-IO. The tableau in (80) illustrates the underapplication of syncope in *fihímna* 'he understood us'.

(80)

Input: /fihim + na/ Base: ['fi.him]	HEADMAX-BA	NO-i]$_\sigma$	MAX-V-IO
a. ☞ [fi.'him.na]		*	
b. ['fhim.na]	*!		*

According to Kager (1999:282), the reason why syncope does apply in the subject forms is that these have no base. A base is in fact defined according to the two criteria in (81).

(81) *Definition of 'base'*

a. The base contains a subset of the grammatical features of the derived form.

b. The base is a free-standing output form – a word.

(ibid.)

While the form *fíhim* 'he understood' is the base of *fihímna* 'he understood us' according to both criteria, it cannot be the base of *fhímna* 'we understood' because it fails (81a). The form *fíhim-* 'understand', on the other hand, cannot be a base because it fails (81b). Since *fhímna* 'we understood' has no base, HEADMAX-BA cannot block the deletion of its first vowel, which is triggered by NO-i]$_\sigma$ (see tableau (82)).

(82)

Input: /fihim + na/ Base: *none*	HEADMAX-BA	NO-i]$_\sigma$	MAX-V-IO
a. [fi.'him.na]		*!	
b. ☞ ['fhim.na]			*

Kiparsky (2000) merges Brame's (1974) and Kager's (1999) analyses in his LPM-OT account of syncope underapplication. While retaining a constraint-based approach, he accepts the notion that subject and object suffixes belong to different morphological strata. The former would belong to the stem level (corresponding to our level 2), the latter to the word level (our level 3). Therefore, while the input to the word level for *fhímna* 'we understood' is *fihímna*, that for *fihímna* 'he understood us' is *fíhim*. Syncope in the latter case is thus blocked by MAX-V́-IO, militating against deletion of stressed vowels, without any need for OO-correspondence (see (83)).

(83)

Word Level	MAX-V́-IO	NO-i]$_\sigma$	MAX-V-IO
Input: ['fi.him]na			
a. ☞ [fi.'him.na]		*	
b. ['fhim.na]	*!		*
Input: [fi.'him.na]			
a. [fi.'him.na]		*!	
b. ☞ ['fhim.na]			*

(from Kiparsky 2000:360)

4.3.2.3 The environment

Not only does syncope generally apply only to monomoraic syllables, but restrictions on the shape of the adjacent syllables are often imposed by the different dialects.

4.3.2.3.1 Domain-initial and medial syncope

As may be seen from the examples in (69) and (70), here repeated together as (84), in AA syncope applies mainly in two environments: a) in a monomoraic syllable which is preceded by another open (but not necessarily monomoraic) syllable (84a),[48] or b) in a domain-initial monomoraic syllable (84b).

(84) (a) *shirib* '(he) drank' ~ *shirbu* '(they) drank'

fihim '(he) understood' ~ *fihmat* '(she) understood'

kibir '(he) grew up' ~ *kibru* '(they) grew up'

SaaHib 'friend (m.)' ~ *SaaHba* 'friend (f.)'

btaaxud '(you m.sg.) take' ~ *btaaxdi* '(you f.sg.) take'

faahim 'understanding (m.sg.)' ~ *faahma* 'understanding (f.sg.)'

[48] In this case, Broselow (1979, 1992) talks about 'high vowel deletion in a 2-sided open syllable', or 'structure-preserving syncope'.

taakul '(you m.sg.) eat' ~ *taaklu* '(you pl.) eat'

saafir 'travel (m.sg.)!' ~ *saafri* 'travel (f.sg.)!'

bitraasil '(you) correspond' ~ *bitraaslu* '(they) correspond'

(b) *muHammad* ~ *mHammad* 'Mohammed'

duru:s ~ *dru:s* 'lessons'

Himaar ~ *Hmaar* 'donkey'

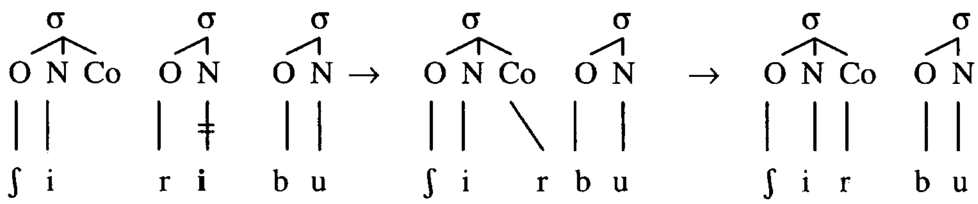
bujuut ~ *bjuut* 'houses'

kilaab ~ *klaab* 'dogs'

kitaab ~ *ktaab* 'book'

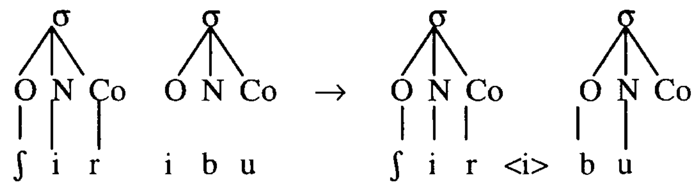
In examples such as those in (84a), the deletion process and its consequences can be straightforwardly explained. When syncope takes place, the initial consonant of the target monomoraic syllable is left floating. However, it is rescued by resyllabification into the coda of the preceding syllable, which is open (see derivation of *shirbu* '(they) drank' in (85)).

(85) *Syncope:* *Resyllabification:*



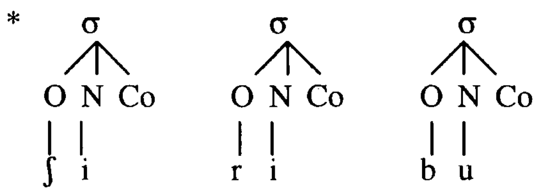
In a directional approach (Itô 1989, Noske 1993), since deletion is part of the syllabification algorithm, there is no need for postulating a resyllabification stage. In this framework, in fact, the second vowel in a word like *shir<i>bu* '(they) drank' cannot be linked to a subsyllabic node in the first place, due to a mismatch between the segmental sequence and the syllabic template. As (86) shows, such a vocoid is matched with an obligatory onset position that it cannot fill, and is thus skipped by the syllabification mechanism. It is therefore left floating, and cannot be phonetically realized.[49]

(86) direction of syllabification ⇒



However, assuming the existence of the syncopated vowel underlingly would obtain the wrong output for dialects in which right-to-left syllabification is postulated (see 4.3.2.3.2 and 5.3). That is, unless specific conditions are stipulated, such a vowel would be preserved (see the erroneous result in (87)); in reality, instead, the target segment is deleted just like in the left-to-right dialects.

(87) ⇐ direction of syllabification



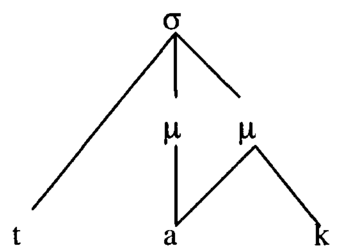
For this reason, most directional accounts assume that all potentially deleted vowels are absent from the UR, and the association of such vowels to words must follow initial syllabification (Broselow 1992:38). The advantage of such an assumption is that it eliminates the need for syncope rules, deriving both syncope and epenthesis from prosodic structuring.

As can be seen from examples such as AA *SaaH<i>ba* 'friend (f.)' or *taak<u>lu* '(you pl.) eat', in most dialects syncope applies even when the vowel of the preceding syllable is long. As was illustrated in 3.3, Arabic dialects adopt different devices in order to eliminate the non-final super-heavy syllable that is thus created. While in varieties like Cairene the long vowel undergoes Closed Syllable Shortening (e.g.

[49] Directional syllabification in AA is assumed to apply from left to right, on the basis of considerations that will be expounded below (but see ch. 5 for problems with this assumption).

/ta:kul-u/ > |ta:klu| > [taklu] '(you pl.) eat'), in AA it is generally preserved. Broselow (1992:14-19), however, suggests that the relevant syllable may preserve its bimoraicity by means of Adjunction-to-Mora, a device creating moras dominating two segments, as in (88).

(88)



In the examples in (84b), syncope affects a word-initial syllable. Consonant clusters are thus formed, many of which actually flout the Sonority Sequencing Principle (e.g. *MuHammad* [**mH**ammed] 'Mohammed').[50] In connected speech, either the first consonant is resyllabified into the coda of a preceding syllable (as in (89a)), or a vowel is epenthesized (as in (89b)).

(89) (a) /abu/ + /**Hu**sayn/ → [abuHse:n] 'Hussein's father'

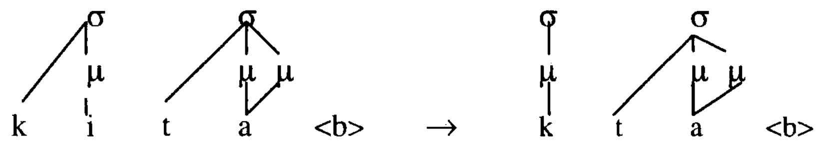
/Tard-i/ + /**ki**bi:r/ → [Tardikbi:r] 'my parcel is big'

(b) /bint/ + /**ki**bi:ra/ → [bintikbi:ra] 'a big girl'

/xudi/ + /l/ + /**ki**ta:b/ → /xudiliktaab/ 'take (f.sg.) the book!'

In absolute domain-initial position, however, resyllabification of the first consonant is not an option. Since, in AA, complex onsets (and codas) are disallowed in all other cases, these clusters must somehow be marked as exceptional. One possibility is to follow Watson (1999a:518-21, 2000:22-3, 2002:76) in assuming that, as a last resort, the mora of the deleted vowel may be preserved and associate with the initial consonant, as in (90).

(90) *kitaab* [kta:b] 'book'



In line with this argument, it has often been pointed out that the forms which have undergone domain-initial syncope are not phonetically identical to hypothetical forms lacking an underlying vowel (e.g. a consonant cluster [b0t] is not the same as a consonant cluster [bt]), i.e. they seem to preserve a 'residue' of the deleted vowel. Harrell (1957:56), with reference to Egyptian Arabic, mentions the presence of a vowel 'microphoneme' |–|, corresponding to a 'voiceless syllable pulse' and alternating with |I| and |U| in forms such as [b–ju:t] 'houses' (from /buju:t/) or [–s_sana:_di] 'this year' (from /ʔis_sana:_di/).[51] These are clearly cases of vowel syncope, and the 'voiceless syllable pulse' seems to correspond to the above mentioned vowel residue.

Moreover, there seems to be a difference in the implementation of the two types of deletion: domain-initial syncope is more optional and more dependent on speech rate and register than medial syncope. This is probably due to the fact that the former creates consonant clusters, while the latter does not.

4.3.2.3.2 Structure-preserving and promiscuous dialects

In some dialects, such as Iraqi, Palestinian and Syrian, syncope applies even when the target syllable is preceded by a closed syllable, as illustrated in (91).

(91) Iraqi: *yilbas* '(he) wears' ~ *yilibsuun* '(they) wear'

---

[50] San9ani Arabic allows syncope to apply to unstressed vowels at the beginning of the PW even when it creates structure-violating clusters of up to three consonants, as in /fitʃih/ > [ftʃih] 'unveiling' (Watson 1999a:517, 2000:20).

[51] Harrell (1957:56) points out that the presence of this 'microphoneme' seems to be dependent on speech rate. However, he mentions a restricted number of lexical items in which a |–| occurs even in slow speech (e.g. /k–rumb/ 'cabbage', /k–wayyis/ 'good (m.sg.)', /ʃ–wayya/ 'little'). Interestingly, in all of these cases the 'pseudo' consonant cluster conforms to Sonority Sequencing.

Syrian: *byəktob* '(he) writes' ~ *byəkətbu* '(they) write'

(Broselow 1992:34)

As is clear from the examples, in these cases epenthesis often applies in order to accommodate the consonant which, not being syllabifiable into the previous syllable, is left stray.[52]

Other dialects, among which Cairene, Sudanese and Makkan Arabic (Broselow 1979, 1992, Majdi and Winston 1994, Zawaydeh 1997), only allow syncope when the preceding syllable is open, i.e. when the resulting string is syllabifiable without any need for epenthesis. AA appears to belong to this latter group, since in this dialect no syncope applies in cases such as those in (92).

(92) *(b)yilbas* '(he) wears' ~ *(b)yilbasu* '(they) wear'

(b)yuktub '(he) writes' ~ *(b)yuktubu* '(they) write'

Broselow (1992:34-5) calls the first type of syncope 'promiscuous', the second 'structure preserving'.

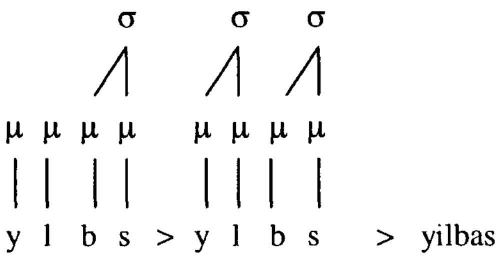
In a directional framework, the divergences between the two groups of dialects are ascribed to a difference in direction of syllabification: while in the 'structure-preserving' dialects syllables would be parsed from left to right, in the 'promiscuous' dialects they would be parsed from right to left.

In most cases, the correct surface forms are thus obtained, as in the examples in (93).[53]

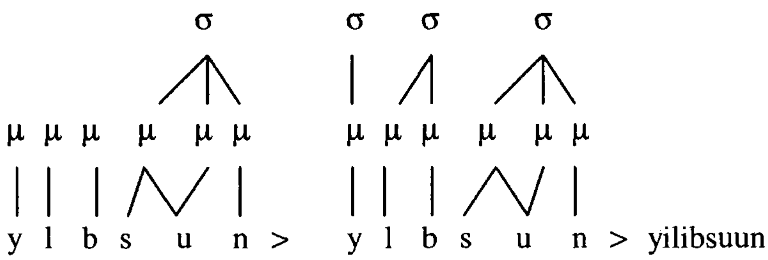
[52] In Tunisian Arabic, not only does syncope apply when the preceding syllable is closed, but the resulting coda cluster is maintained. This creates a non-final superheavy syllable, as in /kahrabitna/ > [kahr.bit.na] 'our electricity' (Majdi and Winston 1994:197). In San9ani Arabic, syncope may even apply to produce a word-medial three-consonant cluster by removing two vowels, as in /yiʃaffiTu:/ > [yiʃfTu:] 'they (m.) make *shafuuT* (Yemeni dish)' (Watson 2000:21).

(93) *Iraqi Arabic (right-to-left):*

(a)



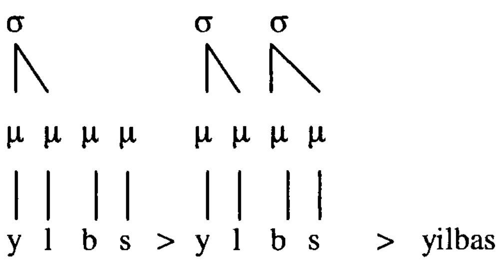
(b)[54]



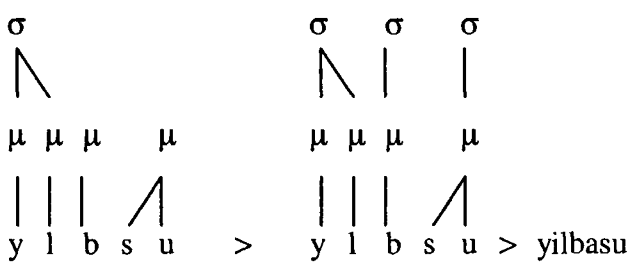
(from Broselow 1992:38)

AA (left-to-right):

(a)



(b)



[53] As was mentioned in 4.3.2.3.1, in this type of directional approach all potentially deleted vowels are absent from the UR.

[54] Broselow (1992) assumes Adjuction-to-Mora even in word-final CVVC syllables, thus avoiding extrasyllabicity of the final consonant.

One first problem, however, is already apparent from these examples. This is how to determine the quality of the epenthetic vowels: if none of them is present in the UR, how is it that a word like *yilbas* '(he) wears' has an [i] in the first syllable and an [a] in the second?

Broselow (1992:39) also shows that some forms cannot be derived by directionality without imposing specific conditions on the prosodic structure of stems and the association of the vowel melody with prosodic templates. Furthermore, the interaction of stress assignment with syncope in certain dialects cannot be accounted for in a templatic framework (ibid.:39-42). The basic problem with the directional account seems to be that it does not draw a distinction between potentially syncopated vowels and epenthetic vowels: in dialects like Syrian Arabic, while the former may be stressed, the latter may not (cf. 5.3).

According to Broselow (1992), the difference between the two sets of dialects lies not so much in whether they allow promiscuous syncope to apply, but in whether they allow nuclear consonants at some level of representation. In the Iraqi-type dialects, the consonant of the deleted syllable, when preceded by a closed syllable and thus left stray, may form a nucleus at some point in the derivation and is only later resyllabified into the rime of an epenthetic syllable. In the Cairene-type dialects, on the other hand, such a consonant cannot be syllabified on its own and the open syllable is immediately restored. Although Broselow (1992) does not assume templatic syllabification and allows syncopated vowels to be present in the UR, her interpretation has something in common with the above directional approach. In order to avoid the conspiracy effect by which, in the 'structure-preserving' dialects, syncope would have to look ahead to see if its output will be syllabifiable, she (ibid.:35-37) assumes that the process applies blindly in both types of dialect to delete monomoraic syllables. By the same token, remedial epenthesis would also apply in both types of dialect. The only differences between the two would be that

a) In Cairene-type dialects, a stray consonant is assigned to the onset of an epenthetic syllable, while in Iraqi-type dialects it is assigned to its rime (cf. 5.3).

b) In Cairene-type dialects, the vowel delinked by syncope escapes Stray Erasure to be resyllabified into the newly formed epenthetic syllable, which is why we obtain AA *yilbas* and not **yilbis* with a default epenthetic vowel.

In conclusion, Broselow's (1992) proposal has at least one drawback in common with the directional account: it does not incorporate the perfectly intuitive and grounded notion that, in the structure-preserving dialects, syncope is blocked exactly in those cases in which its output is not syllabifiable without any need for epenthesis.

Such a drawback, however, cannot even be overcome by the output-oriented OT, otherwise good at expressing conspiracy effects in a natural way. This is because output forms are perfectly syllabifiable not only in Cairene-type dialects, but also in Iraqi-type dialects: in OT, no intermediate stage can be postulated at which a consonant may be syllabic in one dialect group and not in the other. Accounts of syncope (and epenthesis, see 5.3) in Arabic (Mester and Padgett 1994, Wiltshire 1994, Abu Mansour 1995) have therefore been based especially on the notion of alignment, as the one that follows.

In Mashreq dialects such as AA, Cairene, Makkan, Iraqi or Syrian, both MAX-C-IO, which preserves input consonants militating against Consonant Deletion, and *COMPLEX, disallowing complex onsets and codas, are highly ranked. Therefore, they tend to block VD wherever this would create either complex consonant clusters or stray consonants. Another constraint that is virtually inviolate in Arabic is PARSE-MORPH, requiring that every morpheme must be parsed and thus militating against deletion of all segments in a morpheme (Abu Mansour 1995). The differences in output between structure-preserving and promiscuous dialects are instead given by the ranking of the two constraints ALIGN-L (σ, PW) and ALIGN-R (σ, PW). In Iraqi-type varieties, ALIGN-L (σ, PW) will be ranked higher than MAX-V-IO – thus triggering syncope - and DEP-V-IO – thus triggering epenthesis. ALIGN-R (σ, PW) will instead be ranked so low as to be insignificant. On the contrary, ALIGN-R (σ, PW) will be determinant in Cairene-type dialects such as AA. In tableaux (94) and (95), reporting the evaluation of Syrian *byəkətbu* '(they) write' and AA *byuktubu* '(they) write' respectively, the lower-ranking alignment constraint has been left out.

(94) Syrian Arabic

Input: /b-ju-ktub-u/	*COMPLEX[55]	MAX-C-IO	PARSE-MORPH	ALIGN-L (σ, PW)	DEP-V-IO	MAX-V-IO
a. bjək.to.bu				μμ μμμ!		
b. bjək.<tu>bu		*!		μμ		*
c. bjək.t<u>bu	*!			μμ		*
d. bjəkt.<u>bu	*!			μμμ		*
e. ☞ bjə.k[ə]t.<u>bu				μ μμμ	*	*
f. bjək.tob<u>			*!	μμ		*

In tableau (94), the candidate with syncope and epenthesis (i.e. candidate e.) is more harmonic than faithful a. because the sum of the distances of its syllables from the left edge of the PW, calculated in moras, is lower. That is, the first syllable [bjə] of candidate e. is at the edge and thus has a distance of 0 moras, its second syllable [kət] is one mora away from the edge and its third syllable [bu] is three moras away; the total distance is thus of four moras. Candidate a., on the other hand, has a distance of five moras because its second syllable is two moras away from the left edge of the PW and its third syllable is three moras away.

MAX-C-IO, *COMPLEX and PARSE-MORPH, however, still have to be ranked above ALIGN-L (σ, PW) in order to eliminate the suboptimal candidates b., c., d., and f. Their reciprocal ranking, as that between DEP-V-IO and MAX-V-IO, is instead irrelevant to the data at hand.

[55] The problem of whether the initial cluster [bj] violates *COMPLEX is not central to this discussion.

(95) AA

Input: /b-ju-ktub-u/	*COMPLEX	MAX-C-IO	PARSE-MORPH	ALIGN-R (σ, PW)	DEP-V-IO	MAX-V-IO
a. ☞ bjuk.tu.bu				μ μμ		
b. bjuk.<tu>bu		*!		μ		*
c. bjuk.t<u>bu	*!			μ		*
d. bjukt.<u>bu	*!			μ		*
e. bju.k[i]t.<u>bu				μ μμμ!	*	*
f. bjuk.tub<u>			*!	μμ		*

In tableau (95) it is instead the faithful candidate a. that is selected, because it has fewer violations of ALIGN-R (σ, PW) than candidate e. The sum of the distances of its syllables from the right edge of the PW is in fact three moras, while that of e. is four moras.

4.3.3 Other deletion processes

4.3.3.1 Truncation

While in Classical Arabic (CA) (Cantineau 1960:108) and in Modern Standard Arabic (Holes 1995) final short vowels are deleted pre-pausally, truncation is not common in dialects such as AA. Historically, however, word-final vowels underwent a process of deletion in most modern dialects, as can be seen in (96).

(96) AA *katab* 'he wrote' < CA *kataba*

AA *kalbak* 'your dog' < CA *kalbuka*

Some phonologists describe this process as a synchronic one. Abumdas (1985:196-8), for instance, mentions word-final VD (or *apocope*) as applying only pre-pausally in Libyan Arabic. In AA, however, the truncated forms appear in all contexts, so that no alternation takes place; consequently, the process must be considered diachronic.

4.3.3.2 Vowel Elision

As was specified in section 2.3.4 with regard to Glide Formation (GF), it is not easy to find vowel clashes in Arabic because of the full implementation of the Obligatory Onset Principle, which often requires the insertion of a [ʔ] before morpheme-initial vowels. Although clashes do sometimes take place within PUs in varieties like AA, they do not generally induce deletion of one of the two vowels. This is because either at least one of such vowels is high, and then GF is preferred to VD as a remedial strategy, or otherwise the two vowels are identical /a/s. In this latter case, Vowel Contraction may apply, but most often the two vocoids are kept in separate syllables by means of [ʔ]-insertion (97).

(97) *shufna axuuk* [ʃufnaʔaxu:k] or [ʃufna:xu:k] 'we saw your brother'

Layla ajnabiyya [lei̯laʔaʒnabijje] or [lei̯la:ʒnabijje] 'Layla is a foreigner'

sayyaara aaliyya [sajja:raʔa:lijje] or [sajjara:lijje] 'automatic car'

CHAPTER 5: VOCOID INSERTIONS

5.1 Epenthesis

A process of insertion or *epenthesis* is a phonological change by which a segment which is not present in the underlying representation is phonetically realized. Just as in the case of deletions and even more so, the intervention of insertion phenomena can generally be justified on the basis of syllable structure constraints. Both vowels and consonants can be inserted, i.e. the epenthetic sound can occupy different positions inside the syllable. In general, vowels are inserted to resolve consonant clusters, while consonants can be inserted to break up hiatuses, but also to break up consonant clusters (e.g. English *warmth* [wɔ:m***p***θ]) (cf. Picard 1987, Morin 1987). Axininca Campa, for instance, exhibits both Vowel and Consonant Insertion (henceforth VI and CI respectively), as is shown in (1).

(1) /noN-citok-piro-i/[1] → [noncitok***a***piro***t***i] 'I will really hit'

(Itô 1989:237, from Payne 1981)

As we will see in 5.1.2, the vowel [a] is inserted as a syllable peak in order to resolve the ill-formed consonant cluster /kp/, i.e. in order for the stray consonant /k/ to be syllabified as an onset. The consonant [t] is inserted in order to break up the vowel hiatus /oi/, i.e. to give an onset to the otherwise ill-formed syllable $*[i]_\sigma$.[2]

This section is divided into three main subsections, which are organized as follows. Subsection 5.1.1 deals with the insertion of vocoids in onset position, i.e. with so-called Glide Insertion. Vowel epenthesis is instead analyzed in 5.1.2. This subsection first introduces the phenomenon as strictly dependent on syllable structure and delineates the different approaches that have been adopted to deal with it. It then goes on to illustrate various aspects of the implementation of Vowel Insertion, such as the quality of the epenthetic vowel, the location of the epenthetic sites, the relationship

---

[1] In Payne's (1981) notation, <N> stands for a nasal archisegment which is underspecified for place.

[2] As Itô (1989:237) points out, Axininca observes a 'Strict Onset Principle' by which onsetless syllables are not only disfavoured, but completely banned. Therefore, a configuration exhibiting two consecutive syllable peaks is disallowed in this language.

with stress and the interaction with Vowel Deletion. Subsection 5.1.3 introduces an OT interpretation of the different aspects of epenthesis.

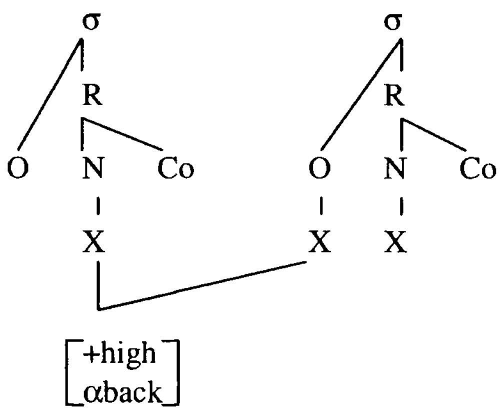
5.1.1 Glide Insertion: vocoids as consonants

As vocoids as a class can often occupy either a nuclear or a marginal position in the syllable, they have the ability to function both as consonant cluster and vowel hiatus breakers. In particular, vocoids are inserted either as syllable peaks or as syllable onsets. In this latter case, they will clearly behave as on-glides, so that the phenomenon is generally referred to as Glide Insertion (GI) (see for example O'Connor 1971:112, Gimson 1989:215-7, Harris 1994:104-5).

As was mentioned in 1.3, onsetless syllables are crosslinguistically dispreferred - although in some languages the ban against them is not absolute - which amounts to saying that hiatuses are phonologically unstable. Therefore, when two nuclear vocoids come into contact, one of three remedial measures may be adopted: Glide Formation (GF, see ch. 2), Glide Deletion (GD, see ch. 4) or CI. GI appears to be a process of the latter type, since on-glides pattern as consonants on functional grounds. The inserted glide, just as an epenthetic consonant, fills in the otherwise empty onset between the two nuclei, which thus become separated by means of a third, non-nuclear element. Rather than as a default consonant, however, the glide can be seen as an extension or a copy of the preceding vowel, whose features spread into the next onset across a syllable boundary, as shown in (2).[3]

[3] It should be noted, though, that a language like English, where resyllabification between words does not take place, distinguishes between expressions such as *my ears* and *my years*, or *two-eyed* and *too wide*: the second phrase in each pair has a longer and stronger glide than the first one (O'Connor 1971:128-9). This difference can probably be explained by the notion that the inserted glide in *my ears* and *two-eyed* is *ambisyllabic*, i.e. it belongs both to the coda of the preceding and to the onset of the following syllable (Sobkowiak 1990:266). In a Romance language like French, on the other hand, j-insertion seems to be a resyllabification phenomenon by which the newly formed glide is exhaustively contained in the onset of the second syllable (cf. Lyche 1979).

(2)



GI generally applies when the first of the two clashing nuclear elements is a high vocoid, since non-high vocoids cannot occupy non-syllabic positions. In Malay, for example, for GI to take place it is necessary that the first vowel is specified as [+high], because a low vowel, being highly sonorous, cannot be copied to form onsets (Durand 1987:96). When the first clashing vowel is not high, a [ʔ] is generally inserted, which gives the perception of the hiatus being maintained (see (3)).

(3)	/peel/	[pe**ʔ**el]	'behaviour'
	/kə-radʒa-an/	[kəradʒa**ʔ**an]	'government'
	/məŋ-gula-i/	[məŋgula**ʔ**i]	'to cause to sweeten'
	/tiap/	[ti**j**ap]	'every'
	/buah/	[bu**w**ah]	'fruit'
	/bantu-an/	[bantu**w**an]	'aid, relief'
	/udʒi-an/	[udʒi**j**an]	'test'

(from Durand 1987:93-4)

It must be pointed out that, in order to trigger GI, the first clashing vocoid does not necessarily occupy a syllable peak: it can be an off-glide which, being placed in the syllable nucleus together with the preceding vowel, cannot function as a hiatus-breaker and thus requires the insertion of an on-glide. In English, for example, GI can take place both after the high vowels [i:] and [u:], as in *the apple* [ði:**j**æpl] or *two of them* [tu:**w**əvðəm], and after the off-glides [i̯] and [u̯], as in *my own* [mai̯**j**əu̯n] or *blowing* [bləu̯**w**iŋ] (see O'Connor 1971:112).

In some languages, moreover, GI may apply after any non-low vowel. This is the case in Dutch, where a word like *vlo* 'flea', though ending in a mid vowel, triggers the intervention of GI (Berendsen and den Os 1987:14).

(4) Hij halt de vlo [w] uit[4] het eten
'He removes the flea from the food'
(ibid.)

However, the quality of the inserted glide always depends on the quality of the first clashing vowel. If such a vowel is front, as in the case of [i] or [e], the inserted glide is generally [j], if it is back (e.g. [o] or [u]) the inserted glide is [w]. This goes to confirm that GI is nothing but the copying of (all or part of) the features of the first vowel into the empty onset slot between the two clashing nuclei.

5.1.1.1 R-Insertion

Although [j] and [w] are by far the most common - and therefore the most commonly inserted – glides, specific languages exhibit hiatus breaking sounds that can be included in this class. R-insertion in non-rhotic varieties of English, for instance, may be seen as a type of GI. Such varieties are characterized by the fact that ⟨r⟩ is only pronounced in syllable-initial position, while nuclear /r/ was dropped in the 18th century (cf. Brown 1988:148): consequently, the well-known phenomena of linking- and intrusive-r ensued (Johansson 1973:56). Linking-r is a process by which, in connected speech, a 'silent' orthographic ⟨r⟩ is pronounced when followed by a vowel, as in (5).

(5)a. tar [tɑ:]→ The ta[ɹ] is still fresh
b. prefer [prɪfə:]→ I prefe[ɹ] orange squash
c. pour [pɔ:]→ pou[ɹ]ing

[4] The diphthong in Dutch <uit> is phonetically realized as [œj], with a mid vowel as its first element. The inserted glide cannot therefore be a copy of the following <u>.

For many non-rhotic speakers, an /r/ is pronounced in an analogous environment (i.e. after [ɑ], [ə] and [ɔ] and before another vowel) even when it is not represented in the spelling, as shown in the examples in (6).

(6)a. spa [spɑ:] → The spa[ɹ] on the mountain.

b. Monica [mɒnɪkə] → Monica[ɹ] is Italian.

c. draw [dɹɔ:] → draw[ɹ]ing.

Apart from the orthographic differences, the two processes are suspiciously similar. Not only the structural change, but also the structural description is actually the same. The only vowels found in word-final position in items ending with an orthographic ⟨r⟩ are [ɔ] (as in *soar*), [ɑ] (as in *far*) and [ə] (as in *fir* or, as the second element of a diphthong, in *dear* and *cure*); intrusive-r occurs exactly after these three sounds (cf. *saw[ɹ] a, spa[ɹ] in, idea[ɹ] of*) (cf. Johansson 1973:54; NV:227; Brown 1988:144). Hence, in synchronic terms, linking- and intrusive-r are one and the same process, and the distinction between them is merely orthographic (cf. Brown 1988:149; Gimson 1989:303). Such a process, which NV (228) call R-Insertion (RI), clearly functions as a hiatus-breaking device by which a dummy segment is inserted between two adjacent vowels in order to avoid a clash of nuclei. Such a segment appears to be ambisyllabic: since resyllabification does not take place in English, there is a sharp phonetic difference between two expressions like *more rice* and *more ice* (Gimson 1989:304). When RI is blocked or avoided out of hypercorrection, other linking phenomena, and specifically the insertion of a [ʔ], take place (Brown 1988:145; Bauer 1984:77).

All these are peculiarities that RI shares with GI. Moreover, the two processes apply in complementary environments: RI never occurs after high vowels, and it is precisely in such a context that GI takes place (Harris 1994:247). Besides, Johansson (1973:60) points out that, since /ɑ:/ and /ɔ:/ can optionally be diphthongised, all the sequences after which [ɹ] is inserted may end in a [ə], which is phonetically closely related to [ɹ]: the alternation [ə] ~ [ɹ] can thus be seen as a sort of vowel ~ glide alternation, similar to [u] ~ [w] or [i] ~ [j]. We can therefore conclude that GI and RI are essentially two aspects of the same hiatus-breaking phenomenon, or better, that RI is a type of GI.

5.1.1.2 Domain of application

In most languages, GI is an optional, postlexical process, taking place in large domains such as the Intonational Phrase (IPh) or the Phonological Utterance (PU). Its application is therefore conditioned by factors such as speech rate[5] and the presence of pauses.

In English, for example, the domain of application of GI (including RI) appears to be the PU (see Bosisio 1998). According to NV's (189) definition of the IPh,[6] the strings in (7) are clearly formed of three, two and three IPh*s* respectively, so that GI each time applies at an IPh juncture.

(7) (a) The kangaroo[w], I imagine, will need a mate.

(b) If you had asked me[j], I would have married you.

(c) The new law[ɹ], as you well know, will damage our economy.

GI, however, does not take place across PUs, as can be seen in the examples in (8), where the process is blocked.[7]

(8) (a) I'd like tea.$]_{U}$ $_{U}[$Andrew's brought some from India.

(b) I think we should go.$]_{U}$ $_{U}[$I'll miss the bus otherwise.

5.1.1.3 Stress assignment

The interaction of GI with stress assignment appears to vary from one language to another. While in English prominence placement does not seem to affect the application of GI, it has been argued that, in other languages, it may block or enhance

[5] At a particularly fast speech rate, GI is often replaced by GF. At a normal rate, however, in languages that do not exhibit postlexical resyllabification GF generally applies only inside the PW. In larger domains, GI substitutes for GF (see Bosisio 1998).

[6] In NV's (189) words, an IPh consists of 'all the PPh*s* in a string that is not attached to the sentence tree at the level of s-structure, or any remaining sequence of adjacent PPh*s* in a root sentence'.

[7] If the two sentences are linked by a strong semantic relation, however, a process of PU-restructuring may take place, by which the two strings are merged into a unique PU (NV 243); in such a case, GI does apply.

the process. For instance, Lyche (1979:328) points out that in French GI can be blocked by an accent on the first (high) vowel, as in (9).

(9)	Un jolí enfant	[ʒo'liã'fã] (*[ʒo'lijã'fã])
	'a pretty child'	
	Il est sí amusant	['siamy'zã] (*['sijamy'zã])
	'he is so funny'	

Similarly, Durand (1987:98) observes that, in Malay, an emphatically stressed initial vowel can trigger the insertion of a [ʔ] rather than other linking phenomena such as GF and GI (Hieke 1984:346). These considerations are not at odds with Wijnen's (1988:200-1) contention that stressed syllables tend to attract consonants, so that GI should be most likely to apply when the second clashing vowel bears prominence. Berendsen and den Os (1987:18), however, advance the contradictory claim that GI in Dutch is highly favoured if the first word is assigned stronger prominence than the second.

5.1.2 Vowel insertion

As was mentioned above, vocoids can also be inserted as syllable peaks. This is a widespread phenomenon that goes under the name of Vowel Insertion (VI), vowel epenthesis or *anaptyxis*, and which generally applies to break up unwanted consonant clusters. In linear phonology, its structural change has been represented as in (10), i.e. as the opposite of VD.

(10) 0 → V

5.1.2.1 A syllable-driven process

It generally applies as a remedial operation in order to avoid phonotactically disallowed strings of consonants which would otherwise be impossible to syllabify in a given language. An unwanted sequence of consonants can be eliminated either by deleting one or more of such consonants by Stray Erasure (e.g. Diola Fogny, Lardil), or by inserting a vocoid between them (e.g. Japanese, Ponapean). In fact, the necessity to avoid complex consonant clusters is by far the most powerful trigger for VI.

In certain dialects of Dutch, for instance, a [ə] is inserted in order to break up clusters formed by a liquid consonant plus a non-coronal obstruent (11).

(11)	pa[rk]	→	pa[rək]	'park'
	he[lp]	→	he[ləp]	'help'
	he[lp]ster	→	he[ləp]ster	'helper (fem.)'

(from NV:78; Nespor 1993:88)

In a linear framework, this process would be represented as in (12).

(12) $0 \rightarrow [ə] \; / \; \begin{bmatrix} +\text{son} \\ -\text{nas} \end{bmatrix} \_\_\_\_ \begin{bmatrix} -\text{son} \\ -\text{cor} \end{bmatrix}$

The examples in (13), however, show that this type of representation misses out on a fundamental consideration, i.e. that VI, like VD, is mostly a syllable-sensitive process.

(13)

par.ken	→	*parəken	'parks'
hel.pen	→	*heləpen	'to help'
mel.ke.rij	→	*meləkerij	'milk farm'

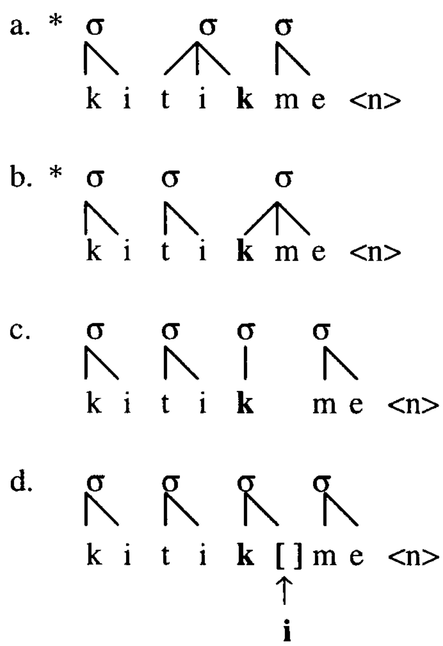
(from NV:78; Nespor 1993:88)

In words like *parken* 'parks', ə-epenthesis does not apply because the two consonants can be syllabified into different syllables. A [ə] is instead inserted when the liquid and the obstruent would have to be syllabified into the same coda, as in (11) (NV:78). Therefore, the process stems from a constraint against tautosyllabic liquid + non-coronal obstruent clusters.

More in general, segmental count and segmental quality considerations are usually insufficient to account for epenthesis processes: one must also refer to the internal structure of syllables.

Language-specific syllable structure constraints account for the implementation of VI in different languages. Ponapean, for instance, has a simple syllable structure characterized by a Coda Filter - restricting codas to the first segment of a geminate or a consonant homorganic to the following segment - and final extrametricality (Itô 1989:229). Therefore, clusters of two non-homorganic consonants are broken up by i-epenthesis, as shown in (14).

(14) /kitik-men/ → kitik**i**men 'rat (indef.)'



(from Itô 1989:230-1)

The two structures in (14a) and (14b) are ruled out by the Coda Filter and by a constraint against complex onsets respectively. Therefore, the [k] is left stray and, because of a parametric choice of the language, rather than undergoing Stray Erasure it is assigned to a syllable of its own (14c). Such a syllable, however, lacks a nuclear position, which is thus inserted as in (14d), and is later filled in by a default epenthetic vowel.

Languages with more complex syllable structure may also require epenthesis in order to break up longer consonant sequences. Northwestern French, for instance, resolves clusters such as /rktr/, /rkm/ or /rkbl/ by means of ə-epenthesis (15a). Two-consonant clusters, on the other hand, do not trigger VI (15b), and neither do three-consonant clusters of the form /rk/ + liquid (15c).

(15)	(a)	Marc Travers (person's name)	→	rkətr
		un turc mort 'a dead Turk'	→	rkəm
		un turc blessé 'a wounded Turk'	→	rkəbl
	(b)	un flic mort 'a dead cop'	→	km
		un sac de sable 'a sack of sand'	→	kd
	(c)	Marc Rainier (person's name)	→	rkr
		Marc Leblanc (person's name)	→	rkl

(from Montreuil 1983)

These patterns are easily accounted for if we assume that Northwestern French disallows complex codas. Moreover, while allowing complex onsets of the form plosive + liquid, it bans onsets with a lower sonority distance (e.g. [km], [kt], [kb]). When a plosive like /k/ is preceded by another consonant, it cannot be syllabified as part of the same coda as the latter, and will therefore be either extrasyllabic (in pre-pausal position) or (re)syllabified into the following onset. When such an onset is already occupied by consonants with which /k/ cannot form a permitted cluster, epenthesis takes place.[8] The examples in (15c) show that, where /k/ can be

[8] The obligatory epenthesis in examples such as those in (i) reinforces the notion that obstruent + liquid + glide clusters are avoided in French, as in Italian (cf. 2.1, 2.2).

(i)	Parc Liais 'Liais Park'	→	rkəlj
	Marc Riochet (person's name)	→	rkərj

(Montreuil 1983:236)

In order to resolve such ill-formed onsets, while Standard French resorts to GF blocking and optional GI (as in *plier* [pli(j)e] 'to bend') a dialect like Norman French resorts to ə-epenthesis between the two consonants (as in *plier* [pəlje]) (ibid.:239).

Montreuil (ibid.:239-242) also shows that ə-insertion applies systematically in the case of CsC clusters. He accounts for this by assuming that in Northwestern French, as in most Romance varieties (see resyllabified into a well-formed onset together with the following liquid, epenthesis does not apply. In (15b), on the other hand, no insertion is required because /k/ is preceded by a vowel, and may thus form a simple coda (e.g. *sac de* [sak.dø] 'sack of').

5.1.2.2 The 'Skeletal Rule' approach

Even among those phonologists who stress the strict dependency of processes such as VD and VI on syllable structure, there is disagreement as to how segmental material is mapped onto syllabic configurations. One approach (Steriade 1982, Clements and Keyser 1983, Archangeli 1984, Levin 1985), which Itô (1989:218) labels as 'Skeletal Rule Theory', formulates insertion processes by referring only to the CV-skeleton, making use of rules like those in (16).

(16) $0 \rightarrow V / C'$ ___

$0 \rightarrow V /$ ___ C'

$0 \rightarrow V / C'$ ___ C'

$0 \rightarrow C / V$ ___ V

Where C' = unsyllabified consonant

(Itô 1989:217)

Alternatively, X-skeletal rules may be adopted, with the proviso that no nuclear vowel can be inserted next to another nuclear vowel, nor can any consonant be inserted next to a stray consonant (Levin 1985:330-1). These have the advantage of collapsing VI and CI into a single process, as in (17).[9]

(17) *X-skeletal epenthesis*

$0 \rightarrow X / X$ ___ X

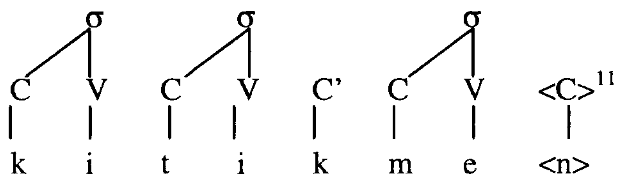
(ibid.:330)

Italian in 2.2 and 4.2), /s/C is not a possible onset.

[9] Itô (1989:238-240), however, argues that such a configuration is inadequate to represent processes such as Axininca Campa 'double epenthesis', which consists of the insertion of a syllable /ta/ whenever a CV-verb is preceded by a word boundary and followed by a consonant-initial suffix (e.g. /na-piro-a:Nci/ > [natapirota:nci] 'to carry well'). In an X-skeletal model, such a process would consist of two consecutive X-insertion operations, the first of which would necessarily produce a VV (naapiro-) or a CC (natpiro-) sequence, thus violating the above-mentioned condition.

This type of framework generally assumes that VI, which is regulated by skeletal rules such as those in (16) and (17), applies when one or more consonants are left stray after initial syllabification. A simple case as that of Ponapean /kitik-men/ > [kitikimen] 'rat (indef.)' is thus analyzed as follows. First, the string undergoes initial syllabification, which leaves the second /k/ unparsed because it cannot be incorporated in the coda of the preceding syllable, nor in the onset of the following one (18a). At this point, the skeletal rule in (18b) applies, inserting a V slot next to the stray consonant (18c). Later syllabification rules build a syllable on the two slots (18d).[10] Finally, the empty slot is automatically assigned default features, which in the case of Ponapean correspond to the segment [i] (18e).

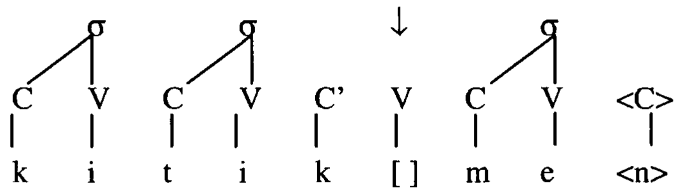
(18) (a)



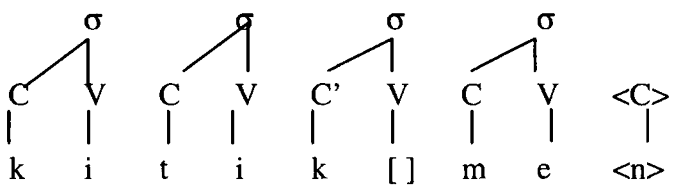
(b) *Ponapean epenthesis rule:*

0 → V / C' ___

(c)

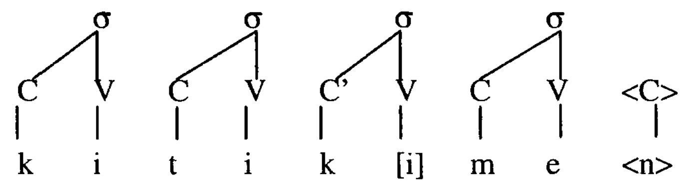


(d)



[10] In Archangeli's (1984:183, 185) model, after the application of epenthesis core syllabification reapplies automatically, reparsing into syllables even those slots which are already syllabified.
[11] The final consonant, as was mentioned in 5.1.2.1, is considered extrametrical.

(e)



The 'Skeletal Rule' approach thus obtains the correct result, but only by making the process at hand look arbitrary by means of a rule such as that in (18b). In Itô's (1989) words, 'the environment of the skeletal rule in (24) [(18b) here] clearly duplicates the syllable structure conditions of Ponapean, which require medial (singly-linked) consonants to be onsets (and not codas).' However, 'since skeletal rules are not intrinsically related to the syllabification mechanism itself, we can imagine a language with the same syllabification parameters as Ponapean but having a different skeletal rule' (ibid.:235), which of course could never exist in reality.

5.1.2.3 Prosodic approaches

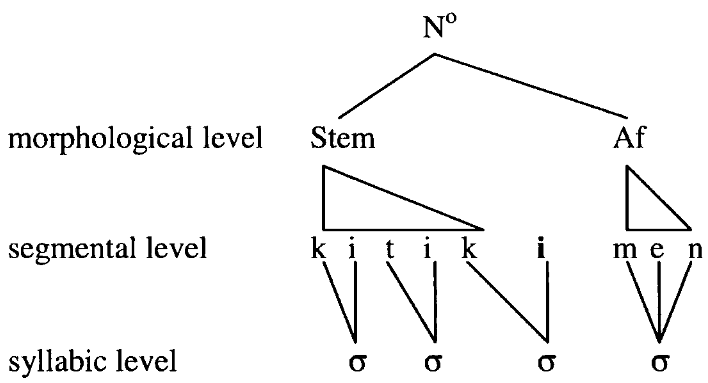
Other approaches treat epenthesis as a prosodic phenomenon stemming directly from syllabification (Selkirk 1981b, Lapointe and Feinstein 1982, Itô 1989, Noske 1993). These models have the advantage of making the relationship between VI and syllabification explicit, thus avoiding 'all diacritic use of "strayness" or "unsyllabified status"' (Itô 1989:218). An account of epenthesis (or of deletion) falls out naturally from the principles and parameters regulating prosodic structure and, more specifically, syllabic structure – such as Sonority Sequencing, the Onset Principle or the Coda Filter (see 1.3). Lapointe and Feinstein (1982:71-2) invoke a 'syllable structure preservation condition', by which rules like VD and VI can make reference to syllable structure, but they cannot create or destroy syllables.[12] They propose that VI, like VD, should operate as part of the 'syllable structure assignment algorithm' (see 4.1).

[12] In a templatic approach assuming post-cyclic syllabification, however, there is no need to stipulate such a condition, since insertions and deletions can be analyzed as a direct result of syllable parsing.

> Under certain universal and language-specific conditions, the syllable assignment algorithm need not attach all of the vowels which appear in the underlying segmental string to a syllable [VD]; and under further conditions, the algorithm may attach certain vowels to a syllable although these vowels are not present in the underlying segmental string [VI].
>
> (ibid.:72)

In their 'dual representation' model (cf. 4.1), contrary to a deleted segment, an 'inserted' segment is one attached to a syllabic node but not to a morphological node, as in (19).

(19) UR: /kitik+men/ 'rat (indef.)'

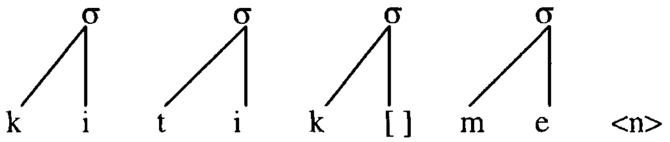


PF: [kitikimen]

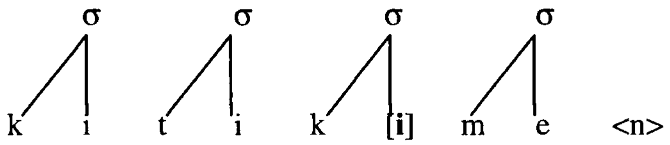
(from Lapointe and Feinstein 1982:80)

As was shown in 5.1.2.2, Skeletal Rule approaches generally allow stray consonants as the output of initial syllabification. Prosodic analyses (cf. Selkirk 1981b, Itô 1989, Broselow 1992), on the other hand, often assume that a syllable is assigned to the extra consonant(s) from the start. That is to say that a nucleus-less degenerate syllable is directly built on the extra consonant during initial syllabification (20a). An epenthetic vowel is later inserted by default rules to give a nucleus to the degenerate syllable (20b).

(20) (a) *Initial syllabification*



(b) *Default vowel insertion*

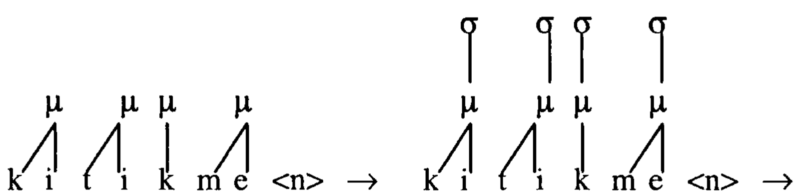


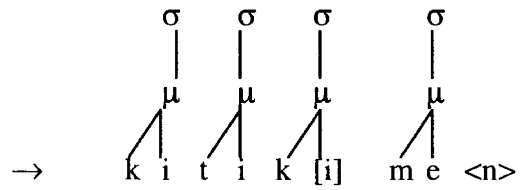
5.1.2.3.1 Templatic syllabification

According to a templatic approach to syllabification (see 1.3), segmental material is inserted in order to satisfy language-specific syllable templates (Itô 1989, Noske 1993). For instance, the syllable template for Ponapean will be (C)V(X), accompanied by final extrametricality and a Coda Filter disallowing coda consonants if they are not place-linked with the following onset. When a sequence such as /kitik-men/ 'rat (indef.)' is syllabified, a mismatch takes place between the segmental string and the syllable template. Since the consonant /k/ does not fit into either the preceding or the following syllable, it is assigned a syllable of its own which, however, has no nucleus. As was said above, such a nucleus is later supplied by default rules.

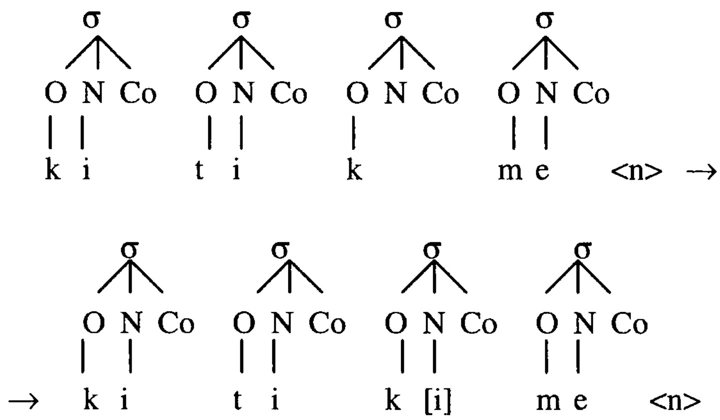
While Itô (1989) postulates that first moraic and then syllabic structure is built directionally on the segmental string (21a), in Noske's (1993) top-down approach each time a syllable is projected its subsyllabic nodes are, too. Therefore, the nuclear position of the degenerate syllable does not need to be inserted, but only filled in with default features (21b).

(21) (a) *Itô's (1989) epenthesis:*





(b) *Noske's (1993) epenthesis:*



5.1.2.4 Quality of the epenthetic vowel

Both the quality of the epenthetic vowel – i.e. the vocoid which is inserted by default whenever epenthesis takes place - and the epenthesis sites – i.e. the exact positions in which the empty syllable peak is found – vary cross-linguistically. The epenthetic vowel, however, tends to be a neutral one, generally [ə] in languages in which this vowel appears underlyingly or as the result of reduction processes (Lapointe and Feinstein 1982:108). In other languages, other – often reduced and less sonorous – vowels are inserted. Alternatively, the epenthetic vocoid may be a copy of a neighbouring vowel. In Ponapean, for instance, the epenthetic vocoid is either a copy of the following vowel (e.g. /ak-dei/ > [ak*e*dei] 'a throwing contest'), or one of the two high vowels [i] and [u] (e.g. [s*i*damp], from 'stamp') (Itô 1989:229).[13]

[13] The melodic character of the inserted segment – i.e. whether it will be a copy of the following vowel or a high vocoid - is determined by morphological factors, probably to do with level-ordering and lexical versus postlexical application of the operation (Itô 1989:229). The choice between [i] and [u], on the other hand, is determined by the phonological environment: the epenthetic vocoid is [u] when near back vowels, [i] in all other cases (ibid.).

The quality of the epenthetic vocoid may also depend on the position in which it is inserted within the syllable. In Temiar, for example, a [ə] is inserted in open syllables, an [e] in closed syllables, as can be seen in (22).[14]

(22)	a.	slɔg	[sə.lɔg]	'sleep/ marry (active, perfective)'
		trakɔ̃w	[tə.ra.kɔ̃w]	'call (causative, simulfactive)'
	b.	sglɔg	[seg.lɔg]	'sleep/ marry (active, continuative)'
		kwkɔ̃w	[kew.kɔ̃w]	'call (active, continuative)'
	c.	srglɔg	[sə.reg.lɔg]	'sleep/ marry (causative, continuative)'
		trwkɔ̃w	[tə.rew.kɔ̃w]	'call (causative, continuative)'

(from Itô 1989:252-253)

5.1.2.5 Epenthesis sites

As was mentioned in 5.1.2.4, the choice between two (or more) possible epenthetic sites varies cross-linguistically, too. Especially in languages exhibiting more complex syllable structures (e.g. lacking a Coda Filter, or allowing for complex onsets and/ or codas), certain stray consonants could potentially be syllabified either as part of an onset or of a coda, and therefore an epenthetic vowel may be inserted either to the left or to the right of such a consonant.

Tigrinya, for instance, exhibits an epenthesis process that operates both word-internally and word-finally (see (23)).

(23) /kalb/ 'dog' + /-n/ "suffixed conjunction" → [kʌlbɨn]

/kalb/ 'dog'→ [kʌlbiː][15]

(from Noske 1993:114)

[14]At least as far as the data in Itô (1989) are concerned, however, an alternative analysis positing a general process of *e*-insertion with subsequent reduction to [ə] in open syllables would be equally valid.

[15] The epenthetic vowel has been lengthened – and consequently tensed - by a process of final lengthening.

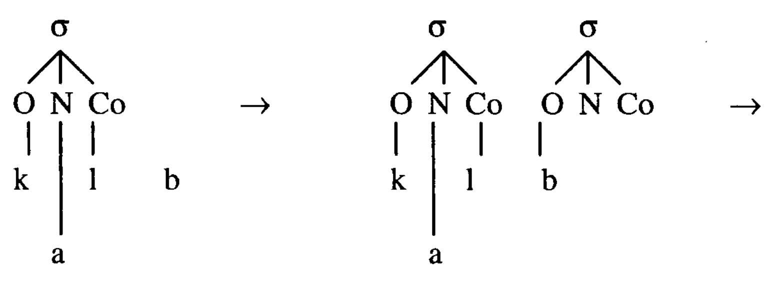
Since the maximal syllable in this language is CVVC, epenthesis intervenes whenever syllabification of a consonant sequence would create a complex onset or coda. While in the first example the only possible insertion site is between [b] and [n], in the second example an epenthetic vowel could in principle be inserted between [l] and [b], yielding the well-formed, but never attested [kʌlɨb]. The choice of the epenthesis site is actually constant throughout the language: in Tigrinya, a CVCC string is always syllabified as CVCCV, never as CVCVC.

This parametrical preference has been interpreted in different ways. One possibility is to assume that Tigrinya is an 'onset language', i.e. a language in which, other things being equal, an extra consonant is syllabified into the onset – rather than the rime - of a degenerate syllable (cf. Selkirk 1981b, Broselow 1992).

An analysis based on directionality of syllabification (Itô 1989, Noske 1993) may also be effective in determining the location of VI, since there is a strict correspondence between right-to-left mapping and Onset Maximisation on the one hand, and left-to-right mapping and Coda Formation on the other (see 1.3).[16]

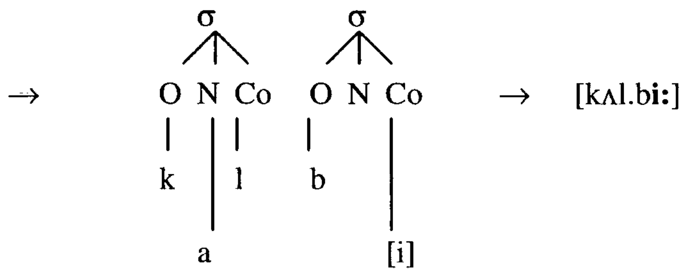
As was mentioned in 4.1, Noske (1993:115) assumes left-to-right syllabification for Tigrinya. This assumption turns out to make the correct prediction as regards the epenthesis sites (cf. 24a); if, on the contrary, syllabification applied from right to left, the surface form for /kalb/ would be *[kʌl**ɨ**b] rather than [kʌlb**iː**] (24b).

(24)[17](a) *Left-to-right syllabification of /kalb/ 'dog'*

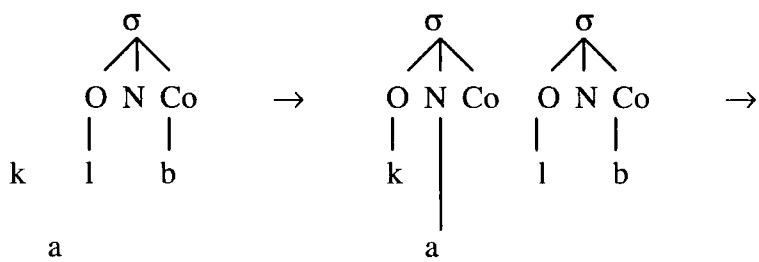


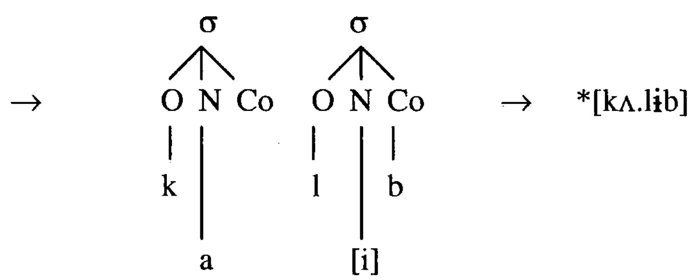
[16] A comparison between the two approaches will be introduced in 5.3, with reference to parametric variation among Arabic dialects (cf. also 4.3).

[17] Vowels and consonants are placed on different tiers because Tigrinya, just like its relative Arabic, is a templatic language.



(b) *Right-to-left syllabification of /kalb/ 'dog'*

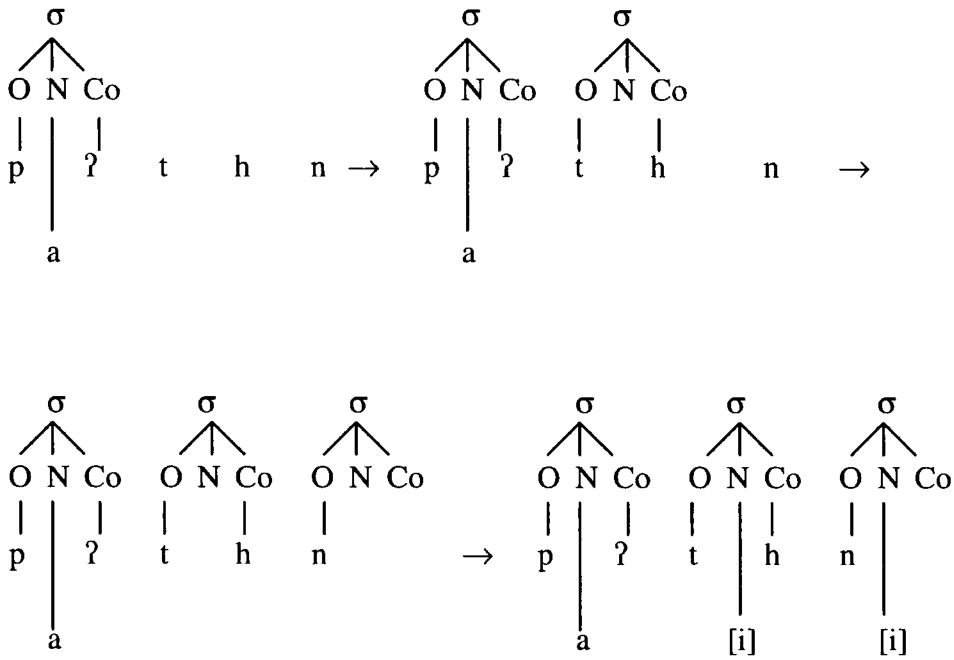




Another templatic language, Yawelmani, has a similar syllabic structure to Tigrinya, in that it lacks a Coda Filter but disallows complex onsets and codas (Archangeli 1984, Noske 1993). Epenthesis thus applies to break up unsyllabifiable consonant clusters, just like in Tigrinya (e.g. /paʔt-hn/[18] > [paʔithin] 'fight (aorist)'). Unlike in the latter language, however, left-to-right syllabification does not predict the correct epenthesis sites (25).

[18] We will assume with Noske (1993), and against Lapointe and Feinstein (1982), that the UR of the aorist suffix is /-hn/ rather than /-hin/.

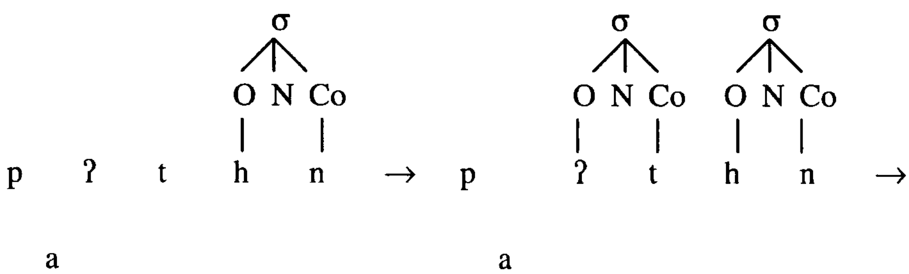
(25) *Left-to-right syllabification of /paʔt-hn/ 'fight (aorist)'*

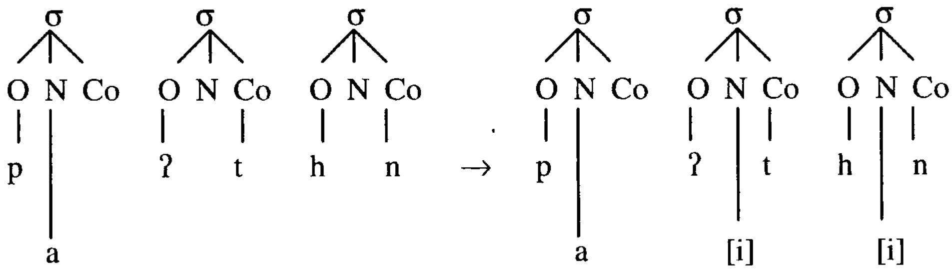


→ *[paʔ.tih.ni]

(26) shows instead that right-to-left mapping, which Noske (1993) postulates to account for deletions, too, obtains the correct surface form for /paʔt-hn/ > [paʔithin] 'fight (aorist)'.

(26) *Right-to-left syllabification of /paʔt-hn/ 'fight (aorist)'*





→ [pa.ʔit.hin]

In conclusion, syllabification and epenthesis in Yawelmani take place exactly as in Tigrinya, only directionality is reversed.

5.1.2.6 Stress assignment

Although the relationship of stress assignment with VI is not as self-evident as that with VD, some epenthesis processes seem to be conditioned by the position of stress. One example is ə-insertion in some varieties of French, where, according to Montreuil (1983:235), the application of the process depends on whether the consonant cluster is immediately followed by a stressed vowel, as in (27).

(27) ours blanc 'polar bear' → [ursə'blã]

BUT

ours blessé 'wounded bear'→ [ursblɛ'se]

Moreover, as will be shown in more detail in 5.3, in many languages epenthetic vowels are never stressed, which seems to indicate that stress assignment applies before epenthesis has taken place.

5.1.2.7 Augmentation

Although most vowel epenthesis processes are straightforwardly related to the internal structure of syllables, there are some cases in which vowel insertion is triggered by other factors. One such case is the augmentation of the aorist of Modern Greek. A morpho-phonological constraint of this language requires that, in the past of perfective verbs, primary word stress fall on the antepenultimate syllable (28a). When a verb root is disyllabic, a word-initial [e] is inserted in order to maintain the syllabic count, thus creating a third syllabic element that can host word prominence (28b).

(28)	a.	[tak'sidepsa]	'(I) travelled'	<	[taksi'devo]	'(I) travel'
		[a'ɣapisa]	'(I) loved'	<	[aɣa'po]	'(I) love'
	b.	['**e**faɣa]	'(I) ate'	<	['fao]	'(I) eat (perfective)'
		['**e**ɣrapsa]	'(I) wrote'	<	['ɣrafo]	'(I) write'

(Nespor 1993:139)

Though they are to some extent morphologically determined, *augments* like these can be considered as epenthetic segments, since they have zero or minimal semantic content.

5.1.2.8 Relationship with Vowel Deletion

Being both often related to syllable structure and phonotactic constraints, it is generally expected that, in a given language, VD and VI should yield identical outputs (see OT analysis of the phenomena). For instance, Montreuil (1983:234) points out that in French there seems to be a correlation between ə-deletion blocking and ə-epenthesis triggering contexts, especially across words. The motivation for the insertion of a [ə] in the examples in (29a) is the same as for the preservation of the [ə] in the examples in (29b).

(29) a. le fisc français [løfisk**ə**frãsɛ]

'the French Inland Revenue'

un ours rouge [ɛ̃nurs**ə**ruʒ]

'a red male bear'

b. un disque français [ɛ̃disk**ə**frãsɛ]

'a French record'

une ourse rouge [ynursəruʒ]

'a red female bear'

(ibid.)

When a segment ~ 0 alternation takes place, therefore, the problem often emerges of whether the alternation must be seen as a deletion or as an insertion. This is all the more so when the alternation takes place within the PW, since when it applies across PW-boundaries the surface forms can be checked against the PWs in isolation.

Orešnik (1972, in Kenstowicz 1994:79-81) argues in favour of *u*-epenthesis rather than deletion in some nominative singular forms of Modern Icelandic. As can be seen in (30), the forms *dagur* 'day (n.sg.)' and *hestur* 'horse (n.sg.)', whose stems end in a consonant, take the suffix *–ur*, while the vowel-ending *bæ-* 'farmhouse' takes a bare *–r* in the nominative singular.

(30)	nom.sg.	dag-ur	hest-ur	bæ-r
	ac.sg.	dag	hest	bæ
		'day'	'horse'	'farmhouse'

(Kenstowicz 1994:79)

It is evident that both interpretations would be perfectly motivated from a phonological point of view: a deletion would avoid a vowel clash in the case of *bæ[0]r*, while an insertion would break up a phonotactically disallowed consonant cluster in cases like *dag[u]r* and *hest[u]r.*

Interaction with two other operations, however, reveals that the process must be one of epenthesis. One such operation is Glide Deletion (GD), by which a stem-final high vocoid is eliminated when followed by a consonant or word boundary (see 31).

(31) /lyfI/ 'medicine'

nom.sg.	lyf-ur
acc.sg.	lyf
gen.sg.	lyfs

dat.pl.	lyfj-um
gen.pl.	lyfj-a

(Kenstowicz 1994:79)

If the UR of the nominative singular ending were [-ur], we would expect GD to be blocked, yielding a surface form **lyfj-ur*. On the contrary, GD does apply, and therefore the UR of the suffix must be [-r].

The second process supporting the epenthesis analysis is *u*-umlaut: when a suffix containing an [u] is added to a stem, an [a] in the preceding syllable surfaces as [ö].

(32)	nom.sg.	hatt-ur	dal-ur	staδ-ur
	acc.sg.	hatt	dal	staδ
	dat.pl.	hött-um	döl-um	stöδ-um
		'hat'	'valley'	'place'

As is shown in (32), the process applies in the case of the dative plural, but it does not in the case of the nominative singular. This goes to show that the latter suffix lacks a vowel at the point in the derivation when *u*-umlaut applies.

From the above evidence, one can thus conclude that the UR for the nominative singular suffix is /-r/, and that such a suffix undergoes *u*-epenthesis when the preceding stem ends in a consonant.

Cross-linguistically, deletion is a more widespread phenomenon than epenthesis (Nespor 1993:82). As can be seen from the examples cited above (*i*-insertion in Yawelmani, *u*-epenthesis in Icelandic, *e*-insertion in Greek, etc.), it is also more limited in its application, and often morphologically conditioned (ibid.:86).

5.1.3 Insertions in OT

The notion that epenthesis is intrinsically linked to syllabification lies at the basis of an OT analysis of the process. This framework, however, does away with language-specific templates and derives the relative strength of prosodic principles from constraint ranking. In OT terms, epenthesis takes place when a segment appears in the output exclusively for phonotactic reasons, and has no correspondent in the input.

Therefore, for insertion to apply, the correspondence constraint DEP-IO (= DEPENDENCY-IO) must be outranked by markedness constraints. DEP-IO is in fact the anti-epenthesis constraint, stating that output segments must have input correspondents. This constraint, just like its mirror image MAX-IO, demands pure segment correspondence, without involving the feature content of the relevant segment (Kager 1999:103).

In *Containment Theory* (PS), FILL ('syllable positions must be filled with underlying segments') is adopted instead of DEP-IO: epenthesis violates FILL in that a syllabic position is phonetically interpreted even though it is not filled in the input. In fact, *Containment Theory* maintains the Templatic Syllabification notion that epenthetic segments correspond to empty syllable positions. This is abandoned by *Correspondence Theory* (McCarthy and Prince 1995), in which epenthetic segments are simply seen as output segments which have no counterparts in the input.

Epenthesis of a consonant into an empty onset is promoted by the ranking ONSET (or NOHIATUS) » DEP-IO. In the specific case of GI, since a high vowel is lengthened to fill in the following onset, the dominated constraint will be DEP-X-IO. Clearly, for GI to apply instead of VD as a repair strategy to break up hiatuses, MAX-V-IO must also be ranked above DEP-X-IO. Tableau (33) exemplifies this ranking on the Malay word /bUah/ > [buwah] 'fruit'.

(33)

Input: /bUah/	ONSET	MAX-V-IO	DEP-X-IO
a. [bu.ah]	*!		
b. ☞ [bu.wah]			*
c. [bu<a>h]		*!	

In languages which do not allow codas, the constraint NO-CODA (or $*C]_{\sigma}$ = 'syllables are open') may trigger VI by dominating DEP-IO. As was shown in 5.1.2, languages that display a Coda Filter (here, Coda Condition) like Ponapean may also require the insertion of a vowel in order to break up heterorganic biconsonantal clusters. In an OT perspective, in cases like these the constraint ranking will be as follows: CODA-COND, IDENT-IO(place) » DEP-IO. That is, it is more important for a coda consonant to share its place of articulation with the following onset consonant, and for output segments to be faithful to their input correspondents in place of articulation, than for all output segments to have a correspondent in the input. A tableau of /kitik-men/ > [kitikimen] 'rat (indef.)' is reported in (34).

(34)

Input: /kitik-men/	CODA-COND	IDENT-IO(place)	DEP-IO
a. [ki.tik.men]	*!		
b. [ki.tip.men]		*!	
c. ☞ [ki.ti.**ki**.men]			*

As was mentioned in 5.1.2, a vowel is often inserted to break up complex onsets and/or codas. In OT, this translates into DEP-IO being dominated by both *COMPLEX (or by one of its components *COMPLEX^{ONS} and *COMPLEX^{COD}) and MAX-C-IO, which prevents the application of Consonant Deletion (CD) as an alternative device (see 4.1). In languages in which coda consonants are tolerated but complex codas are not, the constraint ranking is *COMPLEX^{COD} » DEP-IO » NO-CODA. An example from Tigrinya (/kalb/ > [kʌlbi:] 'dog') is given in (35).

(35)

Input: /kalb/	*COMPLEX^{COD}	DEP-IO	NO-CODA
a. [kalb]	*!		*
b. [ka.l**ɨ**.b**i:**]		**!	
c. ☞ [kal.b**i:**]		*	*

This tableau also shows that the presence of DEP-IO, however low it may be ranked, will always make sure that epenthesis does not overapply, i.e. that no more segments are inserted than are strictly necessary to enforce well-formedness. Besides, unattested types of epenthesis, such as the insertion of a consonant into a marked structural position (e.g. the syllable coda), are automatically ruled out, whatever the constraint ranking (Kager 1999:105).

In other languages, while complex onsets and/or codas are generally allowed (i.e. *COMPLEX is dominated by faithfulness constraints such as DEP-IO and MAX-IO), they are still broken up by epenthesis if they would violate sonority sequencing. In such cases, the constraint ranking will be as follows: SON-SEQ, MAX-IO » DEP-IO » *COMPLEX.

Epenthesis may also apply to satisfy minimal word requirements. As was mentioned in 1.5, the minimal GW often corresponds to a foot, following the requisites that a GW should coincide with a PW (GW=PW) and that every PW should contain at least one foot. If FT-BIN, which requires a foot to be disyllabic or bimoraic, is ranked high and DEP-IO is ranked low enough, a vowel may be inserted into a monosyllabic word to form an extra syllable and render the foot, and thus the word, disyllabic. Therefore, epenthesis enforces the word minimum when the ranking of these three constraints is as follows:

GW=PW, FT-BIN » DEP-IO

Since epenthetic segments are not present in the input, they cannot be subject to faithfulness (i.e. input identity) constraints, so that their quality is entirely determined by markedness factors. Therefore, they either exhibit maximally unmarked feature values or are influenced by the feature make-up of neighbouring segments. Typically unmarked feature values are [-low], hence [ə], [i] and [u] are often inserted, [-round], hence [ə] and [i] are inserted more often than [u], but also [-front], hence [ə] is inserted more readily than [i]. In fact, [ə], unmarked vowel *par excellence*, is generally inserted whenever it is available from the phonetic inventory of a language, unless contextual factors intervene. Similar considerations apply to consonants: the most commonly inserted are the least marked cross-linguistically, such as [ʔ], which lacks supralaryngeal specifications, [t], which exhibits the unmarked value for place of articulation ([+cor]) and for voice in stops ([-voice]), or even [j], as the 'consonantal' counterpart of the vowel [i].

As for the influence of neighbouring segments, this can be expressed in terms of contextual markedness constraints. Since input faithfulness, which tends to counteract assimilation, does not affect epenthetic segments, such segments are open to the influence of contextual markedness. The quality of epenthetic segments thus totally depends on the ranking of these two types of constraints: if context-free markedness prevails, the epenthetic segment is minimally marked; if contextual markedness prevails, the epenthetic segment is contextually coloured. In reality, things are more complex than this, since single context-free and contextual markedness constraints may interact independently.

The fact that epenthetic vowels often escape stress assignment is accounted for by high ranking of the constraint HEADDEP-IO; this dictates that 'every vowel in the output's prosodic head has a correspondent in the input'.

In order to predict the location of epenthesis sites, OT often resorts to gradient alignment constraints such as ALIGN-R (σ, PW) – by which the right edge of a syllable must be aligned with the right edge of a PW - and ALIGN-L (σ, PW) - by which the left edge of a syllable must be aligned with the left edge of a PW (Mester and Padgett 1994, Wiltshire 1994). As was mentioned in 4.1 and 4.3, violations of such constraints are measured by the sum of the number of moras that separate each misaligned syllable from the relevant edge. Such a number changes according to the way in which segments are syllabified, which in its turn may depend on the location of the epenthetic vowel. In a language like Yawelmani, for instance, ALIGN-L (σ, PW) must outrank ALIGN-R (σ, PW), as the tableau of /paʔt-hn/ > [paʔithin] 'fight (aorist)' in (36) illustrates.

(36)

Input: /paʔt-hn/	*COMPLEX	ALIGN-L (σ, PW)	ALIGN-R (σ, PW)
a. [paʔthn]	*!		
b. ☞ [pa.ʔit.hin]		μ μμμ	μμ μμμμ
c. [paʔ.tih.ni]		μμ μμμ!μ	μ μμμ

In tableau (36), the optimal candidate b. receives fewer violations of ALIGN-L (σ, PW) than suboptimal c., because its second syllable is one mora away from the left edge of the PW and its third syllable is three moras away, for a total of four moras. The second syllable of candidate c. is instead two moras away from the relevant edge, and its third syllable is four moras away, for a total of six moras. If the ranking between ALIGN-R (σ, PW) and ALIGN-L (σ, PW) were reversed, c. would be the winning candidate, because it has fewer violations of the former constraint than b. has.

As will be shown in 5.3, however, in certain varieties the outputs of syncope and epenthesis would require opposite rankings of the alignment constraints, putting their validity into question (Zawaydeh 1997).

The location of epenthesis may also be determined by correspondence constraints, such as ANCHOR-IO, by which any segment at the left/right periphery of the output grammatical constituent has a correspondent at the left/right periphery of the input grammatical constituent. As was mentioned in 4.1, this latter constraint has the effect of penalizing epenthesis (or deletion) at constituents' edges. Another constraint that may determine the epenthesis sites in a given language is CONTIG-IO, which militates against (morpheme-)medial epenthesis (or deletion) by demanding that input-output mappings involve continuous substrings.

5.2 Insertions in Italian

In Italian, insertions of both vowels and consonants are not common and evenly distributed phenomena: they tend to be limited in their application by lexical, diatopic and diastratic factors. After briefly introducing the lexically conditioned phenomenon of CI, subsection 5.2.1 deals with GI in Italian. A more articulated second subsection analyzes the issue of vowel epenthesis in its various forms. It first shows how, in certain varieties of Italian, disallowed coda consonants and consonant clusters may be resolved by means of VI, but also how this is by no means the only strategy adopted in the language. It then goes on to expand on the already mentioned phenomenon of 'pseudo-epenthesis' (see 4.2), comparing Burzio's (1987) skeletal approach with a directional one. As usual, OT accounts of the various processes are also given.

5.2.1 Consonant and Glide Insertion

In Italian, both VI and CI apply. The latter process is a sporadic phenomenon, generally restricted to specific words and almost always fully lexicalized. The epenthetic consonant is often [d], as can be seen in (37a), but there are a few isolated cases in which a different consonant is inserted (37b).

(37) a. cittá 'city' + -ino → citta[d]ino 'citizen' (*cittaino)
a 'in' + Istanbul → a[d] Istanbul ~ a Istanbul
Roma 'Rome' + e 'and' + Atene 'Athens' → Roma e[d] Atene ~ Roma e Atene

b. caffé 'coffee' + -iera → caffe[tt]iera 'coffee-pot' (*caffeiera)
papá 'dad' + -ino → papa[r]ino 'daddy' (*papaino)
Forlí (an Italian city) + -ese → forli[v]ese 'from Forlí' (*forliese)

(from Nespor 1993:89)

The examples in (37) also show that inside the PW the process is obligatory, while in postlexical domains it is optional.

GI, on the other hand, is a fully productive process in Italian, even though it is optional and less common than in Germanic languages like English or Dutch. As in the latter languages, it generally applies when GF fails to take place. However, since its domain of application in Italian is the PU, GF replaces GI in many postlexical contexts (see 2.2).

As was shown in 5.1, in many languages hiatuses are absolutely banned, so that, when GF cannot apply, a process of CI – be it in the form of GI or of the insertion of another consonant, most commonly [ʔ] – obligatorily takes place. In Italian, on the other hand, sequences of syllable nuclei, while clearly disfavoured, are allowed in certain circumstances, and even preferred to some repair strategies. ʔ-insertion, for instance, is very rare, since this language lacks glottal sounds in its phonetic inventory. Although more common, GI is also not obligatory as a substitute for GF. In the sentences in (38), this latter process is blocked: in (38a), (b) and (c) because of a main stress falling on the high clashing vocoid, in (38d) because of a preceding obstruent + liquid cluster [tr] (cf. 2.2).

(38) a. Mercoledí [j/0] ho fatto una torta.
'On Wednesday I baked a cake'

b. Cosí [j/0] Alberto ha perso di nuovo.
'So Alberto has lost again'

c. Ho visto un caribú [w/0] albino.
'I have seen an albino caribou'

d. I maestri [j/0] hanno deciso di scioperare.
'Schoolteachers have decided to go on strike'

In each of these examples, the insertion of a glide between the clashing vowels often fails to occur: apparently, the ban against hiatuses is less strong in Italian than in other languages.

In an OT framework, the trigger constraint for GI, as for GF, is NoHiatus (or Onset). In order for GI to take place, as was explained in 5.1.3, this constraint must outrank Dep-X-IO, which would otherwise prevent the lengthening of the high vowel into the following onset. Constraints blocking VD – such as Max-V-IO - and GF – such as Pk-Max-IO, militating against GF from stressed vowels or *OLG, disallowing obstruent + liquid + glide clusters – must also dominate Dep-X-IO. However, since GI is an optional process and hiatuses are often preserved, free ranking between NoHiatus and Dep-X-IO must be assumed. Tableaux for *cosí Alberto* 'so Alberto' ([ko.'zi.**j**al.'bɛr.to] or [ko.'zi.al.'bɛr.to]) according to the two subhierarchies are given in (39a) and (39b).

(39) (a) Subhierarchy 1 (NoHiatus » Dep-X-IO)

Input: /ko'sI al'bɛrto/	Pk-Max-IO	Max-V-IO	NoHiatus	Dep-X-IO
a. [ko.'zi.al.'bɛr.to]			*!	
b. [ko.'zjal.'bɛr.to	*!			
c. [ko.'zil.'bɛr.to]		*!		
d. ☞ [ko.'zi.**j**al.'bɛr.to]				*

(b) Subhierarchy 2 (Dep-X-IO » NoHiatus)

Input: /ko'sI al'bɛrto/	Pk-Max-IO	Max-V-IO	Dep-X-IO	NoHiatus
a. ☞ [ko.'zi.al.'bɛr.to]				*
b. [ko.'zjal.'bɛr.to	*!			
c. [ko.'zil.'bɛr.to]		*!		
d. [ko.'zi.jal.'bɛr.to]			*!	

5.2.2 Vowel Insertion

As described in section 1.3, Italian syllable structure is rather simple. While complex onsets are allowed – although with strong phonotactic restrictions - not only complex, but also simple codas are generally disfavoured. This is due to a modified Coda Filter by which only the first segment of a geminate, a sonorant consonant or an [s] are allowed in such a position, and even these are often avoided word-finally. In fact, Italian appears to lack final extrametricality. The Onset Maximization Principle is implemented whenever possible, so that when a string of consonants is found, only those which cannot be syllabified into the onset of one syllable will be left for syllabification into the coda of the previous one (e.g. *con.tro* 'against', *se.pol.cro* 'sepulchre'). As was shown in section 5.1, when a consonant cannot be syllabified either in an onset or in the preceding coda, it is left floating, and therefore cannot be licensed. Since prosodic licensing is obligatory, one of two options will apply: either the stray consonant is assigned to an extra syllable whose nucleus is filled in by an epenthetic vowel (VI), or it is not realized phonetically (CD). As we will see, both processes may apply in Italian.

5.2.2.1 Word-final consonants

In this language, however, the application of VI is even more sporadic and variable than that of CI and GI. According to Lapointe and Feinstein (1982), languages that require a simpler syllable structure appear to resort more often to deletion than to epenthesis when resolving complex consonant clusters. Furthermore, complex clusters are rarely found in Italian even underlyingly, since the overwhelming majority of lexical items in this language end in a vowel. Unwanted coda consonants can generally be found only in loanwords such as *tram*, *jet* or *top*. Even in such cases, however, VI is by no means an evenly widespread phenomenon: it is generally confined to specific regions and social classes, and it is never obligatory even where it is most favoured. Moreover, its implementation varies widely according to the place, the speaker and even the specific situation.

For instance, in the Central and Southern varieties of Italian VI is often accompanied by doubling of the triggering consonant (e.g. *top* → [toppe]). That is, bimoraicity of the stressed syllable – which in these varieties, just like in Standard Italian (SI), is obligatory (cf. 3.2) - is maintained by geminating the consonant rather than by lengthening the preceding vowel. On the contrary, in Northern Italian – where the phenomenon is anyway less common and more restricted in scope – the consonant is not doubled.[19] Besides, the quality of the epenthetic vowel may vary considerably from one place to another, sometimes yielding a [ə], sometimes an [e], sometimes an [i]. This seems to substantiate Burzio's (1987) claim that (Standard) Italian does not

[19] As was mentioned in 3.2, in the Northern varieties of Italian geminate consonants are generally dispreferred or even, for certain speakers, completely banned. Moreover, stressed syllables are not always bimoraic.

have a specific epenthetic vowel, and that it fills in pseudo-epenthetic gaps only when a floating vocoid is present (see 4.2 and 5.2.2.3).

It must also be kept in mind that VI is by no means the only available device to resolve an unwanted consonant cluster or an unsyllabifiable final consonant: CD is a widespread alternative. NV (143) point out that the type of readjustment adopted varies depending on the region, on the structure of the last syllable and even on the type of consonant involved. In Tuscany, for instance, a word-final [l] is usually deleted, as in (40), while an [m] triggers epenthesis (see (41)). According to the number of syllables in the word, the [m] can remain simple, as in (41b), or be doubled, as in (41a).

(40) a. alcol → ['alko] 'alcohol'
b. würstel → ['vuste] 'wurstel sausage'

(41) a. tram → ['tram:**e**] 'tram'
b. reclam → [re'kla:m**e**] 'advertisement'

It should be noted, however, that the difference between the forms in (40) and those in (41) lies not only in the quality of the word-final consonant, but also in the distribution of PW-stress. While in (41) main stress falls on the vowel immediately preceding the triggering consonant, in (40) it falls on the previous syllable. The reason why a word-final stress favours epenthesis is probably related to a preference for binary feet. Since feet in Italian are left-headed, an epenthetic syllable in words like those in (40) would form a marked trisyllabic foot together with the two preceding syllables. In *tram* ['tram:**e**] 'tram' and *reclam* [re'kla:m**e**] 'advertisement', on the other hand, epenthesis yields unmarked disyllabic feet. This observation is confirmed by Lepschy and Lepschy (1977:68), who point out that, in Florentine Italian, consonant endings tend to undergo *e*-epenthesis especially if the final syllable is stressed. Two words like *gas* 'gas' and *lapis* 'pen', though ending in the same consonant, yield the surface forms ['gass**e**] and ['la:pis] (or even ['la:pi]) respectively.

In a constraint-based approach, word-final consonants are disallowed by the markedness constraint *FINAL-C, by which all words end in a vowel. This constraint must clearly be kept distinct from CODA-COND(mod), which corresponds to the Italian modified Coda Filter, since this latter does allow liquids, [s] and place-linked consonants in coda position. In order for CD and VI to take place, *FINAL-C must outrank both MAX-C-IO – requiring input consonants to have correspondents in the output - and DEP-V-IO, disallowing output vowels without a correspondent in the input. The ranking between these two constraints is MAX-C-IO » DEP-V-IO since, other things being equal, VI is preferred over CD. However, since FT-BIN (requiring feet to be binary) dominates MAX-C-IO, CD applies when VI would produce a trisyllabic foot, as in the case of *alcol* ['alko] 'alcohol' (cf. tableau (42)).

(42)

Input: /alkol/	*FINAL-C	FT-BIN	MAX-C-IO	EVEN-TROCHEE	DEP-V-IO
a. ['al.kol] F	*!				
b. ['al.ko.**le**] F		*!			*
c. ☞ ['al.ko]F			*	*	

For comparison's sake, the evaluation of the candidates for *tram* ['tram**me**] 'tram' is reported in tableau (43). This time, the candidate with epenthesis does not violate FT-BIN, and is therefore selected.

(43)

Input: /tram/	*FINAL-C	FT-BIN	MAX-C-IO	EVEN-TROCHEE	DEP-V-IO
a. ['tram] F	*!				
b. ☞ ['tram.me]F				*	*
c. ['tra]F		*!	*	*	
d. ['tra:]F			*!		

It is clear from these tableaux that, since both optimal candidates violate EVEN-TROCHEE, this must be ranked at least below *FINAL-C and MAX-C-IO.

It must be pointed out, however, that this is the constraint ranking for Tuscan Italian. In SI, where neither VI nor CD take place, faithfulness constraints such as DEP-V-IO and MAX-C-IO must outrank *FINAL-C.

5.2.2.2 Disallowed consonant clusters

As for phonotactically disallowed consonant clusters, these are found especially in learned and technical terms and loan-words from Greek and Latin. They are generally formed by an obstruent plus another consonant, and can be found both word-internally - as in *tecnico* ['tɛ**kn**iko] 'technical' or *atmosfera* [a**tm**os'fɛ:ra] 'atmosphere' – and word-initially, as in *pneumatico* [pneu̯'ma:tiko] 'pneumatic' or *xilofono* [ksi'lɔ:fono] 'xylophone'. Here again, the intervention of VI is optional, variable and unevenly distributed, though always motivated by syllable structure constraints. When the unwanted cluster is word-medial, its first consonant cannot be syllabified in the coda of the preceding syllable because of the Coda Filter banning plosives from such a position; it cannot be syllabified into the following onset, either, because the sonority differential between it and the next sound is too low. Therefore, the relevant consonant is left stray. In colloquial Italian, its condition may be resolved either by total assimilation to the other consonant to create a geminate (e.g. *atmosfera* [ammos'fɛ:ra], typical of Tuscany, Umbria, Marche, Lazio, Sicily), or by the insertion of a vowel between the two consonants (Serianni 1989:35). Such a vowel is typically a [ə] in the continental South (e.g. *atmosfera* [atəmos'fɛ:ra]), a copy of an adjacent vowel in the North-West, where, contrary to the continental South, the sound [ə] is not available from the dialectal substratum (e.g. *atmosfera* [atomos'fe:ra], *tecnico* ['tekeniko]).

Word-initial 'learned' clusters tend to be avoided because of the same syllable structure and phonotactic constraints: if maintained, they constitute ill-formed onsets or disallowed codas. Therefore, they are generally broken up by means of vowel epenthesis, yielding forms such as *psicologo* [pisikɔ:logo] 'psychologist'. It should be borne in mind, however, that all of these forms are highly stigmatized from a social and regional point of view.

In an OT perspective, the ban on word-medial plosive + nasal and plosive + fricative clusters can be expressed by high ranking of two constraints. One of these is CODA-COND(mod), which disallows, among others, singly-linked coda plosives. The other constraint, which prevents the relevant clusters from being exhaustively syllabified in an onset, will be called SON-DIS.[20] This is a cover constraint which includes other more specific prohibitions, such as *PN and *PF, disallowing plosive + nasal and plosive + fricative onset clusters, respectively. These constraints are clearly related to sonority distance limitations, but a unified distance limit cannot be quantified because

[20] This constraint is not to be confused with SON-SEQ, which only refers to the increase in sonority of onset and coda towards the nucleus. Since the relevant clusters regularly rise in sonority, SON-SEQ is not violated.

onset clusters such as fricative + liquid, which have an analogous – if not inferior - distance to plosive + nasal, are allowed in Italian (cf. ***flusso*** 'flow').

In varieties of Italian in which assimilation is preferred over VI, CODA-COND(mod) and SON-DIS outrank IDENT-IO, which requires that correspondent segments have identical values for all features, thus militating against assimilation. In order to prevent VI, IDENT-IO must also be dominated by DEP-V-IO (cf. tableau (44)).

(44)

Input: /atmosfɛra/	CODA-COND(mod)	SON-DIS	DEP-V-IO	IDENT-IO
a. [at.mos.fɛ:.ra]	*!			
b. [a.tmos.fɛ:.ra]		*!		
c. [a.tə.mos.fɛ:.ra]			*!	
d. ☞ [am.mos.fɛ:.ra]				*

Those varieties which adopt VI, rather than assimilation, as a remedial strategy, will exhibit the ranking IDENT-IO » DEP-V-IO (cf. tableau (45)).[21]

(45)

Input: /atmosfɛra/	CODA-COND(mod)	SON-DIS	IDENT-IO	DEP-V-IO
a. [at.mos.fɛ:.ra]	*!			
b. [a.tmos.fɛ:.ra]		*!		
c. ☞ [a.tə.mos.fɛ:.ra]				*
d. [am.mos.fɛ:.ra]			*!	

In the case of domain-initial clusters, assimilation is always ruled out as a remedial strategy, because it would bring no gains to compensate for the violation of high-ranked constraints such as $*[...C_1C_1 ...]_\sigma$, disallowing tautosyllabic geminates. In order for the geminate not to be tautosyllabic, a vowel should be inserted before it, adding a violation of DEP-V-IO to that of IDENT-IO. As the tableaux of *psicologo* [pisikɔ:logo] 'psychologist' in (46) show, the candidate with epenthesis is always one violation ahead, whatever the ranking between DEP-V-IO and IDENT-IO.

[21] Since CD never applies, it is assumed that MAX-C-IO dominates both these constraints in both dialect types.

(46) (a)

Input: /psikɔlogo/	*[…C₁C₁ …]σ	CODA-COND(mod)	SON-DIS	DEP-V-IO	IDENT-IO
a. [psi.kɔ:.lo.go]			*!		
b. ☞ [**pi**.si.kɔ:.lo.go]				*	
c. [ssi.kɔ:.lo.go]	*!				*
d. [**is**.si.kɔ:.lo.go]				*	*!
e. [**ip**.si.kɔ:.lo.go		*!		*	

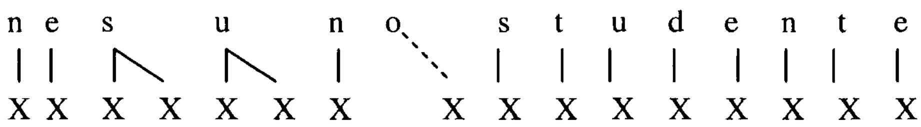
(b)

Input: /psikɔlogo/	*[…C₁C₁ …]σ	CODA-COND(mod)	SON-DIS	IDENT-IO	DEP-V-IO
a. [psi.kɔ:.lo.go]			*!		
b. ☞ [**pi**.si.kɔ:.lo.go]					*
c. [ssi.kɔ:.lo.go]	*!			*	
d. [**is**.si.kɔ:.lo.go]				*!	*
e. [**ip**.si.kɔ:.lo.go		*!			*

5.2.2.3 Burzio's pseudo-epenthesis

As was mentioned in section 4.2, Burzio (1987) talks about 'pseudo-epenthesis', rather than 'epenthesis' proper, when dealing with certain forms of SI. Such forms are precisely those exhibiting initial clusters which cannot be exhaustively included into one onset, such as /s/+C, /ps/, /ks/, /pn/, /pt/, /kt/ and the consonants which are always double (/ɲ/, /ʃ/, /ʎ/, /ts/ /dz/ and /j/). Since, contrary to other varieties of Italian, the Standard variety does not exhibit an epenthetic vowel, the process is limited to the insertion of a time unit, which Burzio (ibid.) calls 'augmentation'. Such a time unit may be filled with segmental material only when this is available from another word, i.e. when it is left floating as a result of 'prosodic reduction' (Burzio 1987:6). The latter phenomenon, which was dealt with in more detail in section 4.2, applies to masculine noun specifiers such as *lo* 'the' and its compounds, *uno* 'a, one', *quello* 'that', *bello* 'nice, handsome, beautiful', *alcuno* 'any, some', *nessuno* 'no', *buono* 'good'. These lose their final vowel in all contexts (e.g. *un buon amico* 'a good friend (m.)', *un buon pranzo* 'a good lunch') except when the following word begins with one of the above-mentioned clusters (e.g. *un buono psicologo* 'a good psychologist (m.)'). According to Burzio (ibid.) this can be satisfactorily explained by postulating that the final /o/, which is left floating by 'prosodic reduction', goes to fill in the timing slot which is inserted at the beginning of the cluster-initial words, thus being phonetically realized. This is called ‘pseudo-epenthesis’. The two combined processes are exemplified in (47), on the phrase *nessuno studente* ‘no student’.

(47) a. PROSODIC REDUCTION b. PSEUDO-EPENTHESIS



(from Burzio 1987:6)

Just as the floating vowel will not surface if it cannot be linked to an empty timing slot, pseudo-epenthesis ‘will only have any visible effect when there is a floating vowel to be associated with the pseudo-epenthetic time unit’ (ibid.).

The plausibility of such an analysis becomes clear when Italian is compared to Spanish, a related language that does have an epenthetic vowel [e]. Contrary to Italian *studente* ‘student’, a word like Spanish *estudiante* ‘student’ regularly undergoes epenthesis in all contexts. In fact, *e*-insertion in Spanish takes place obligatorily before a word-initial /s/C cluster, since such a cluster is absolutely banned from onset positions in this language.[22] Even if the process is limited to this type of clusters, it is unexceptional and fully productive, so that loanwords such as *estrés* ‘stress’ *estalinismo* ‘Stalinism’ and *espray* ‘spray’ must undergo it (Nespor 1993:86). In SI, on the other hand, /s/C clusters - like all other above-mentioned clusters - are only disfavoured, and surface wherever a preceding vowel is not available (i.e. post-consonantally, in isolation and in post-pausal position). The two Romance languages are compared in (48).

(48) *Spanish* *Italian*

[e]scuela scuola ‘school’

[22] It must be kept in mind that onsets formed by /s/+plosive clusters like /st/ are actually ill-formed cross-linguistically, since they violate sonority sequencing, by which onset clusters must increase in sonority. According to some phonologists (see for example NV), an /s/ followed by a plosive is always syllabified in a coda of a syllable whose nucleus may be empty when no vocoid is available to fill it.

[e]speranza	speranza	'hope'
[e]star	stare	'to stay/ be'
[e]smalte	smalto	'enamel'
[e]sdrújulo	sdrucciolo	'proparoxytone'

(from Nespor 1993:86)

5.2.2.3.1 The plural

As was shown in section 4.2, Burzio (1987:7-10) advances the same proposal to deal with alternations in the plural forms, such as *belli* [bɛlli] ~ *bei* [bɛi] ~ *begli* [bɛʎʎi] 'nice, beautiful, good-looking (m.pl.)' or *quelli* [kwelli] ~ *quei* [kwei] ~ *quegli* [kweʎʎi] 'those (m.)'. In each case, the first form appears in isolation (49a), the third before vowels and the above-mentioned geminate consonants and consonant clusters (49c) and the second before all other consonants and clusters (49b).

(49) a. Sono belli

'(They) are beautiful (m.)'

Voglio quelli

'(I) want those (m.)'

b. Bei gatti

'Beautiful cats'

Quei criceti

'Those hamsters'

c. Begli uccelli

'Beautiful birds'

Quegli psicofarmaci

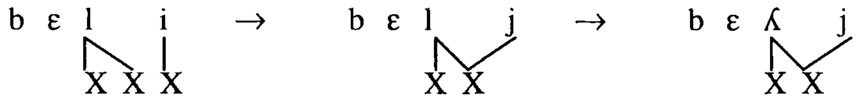
'Those psychotropic drugs'

Negli gnocchi [ɲɲɔkki]

'In the dumplings'

According to Burzio (ibid.:7), prosodic reduction would apply to these forms as well, deleting the final time unit as in the case of *bello/ quello.* As a result of this process, the /i/ of the plural ending, rather than undergoing complete delinking and being left floating, would lose its syllabicity and become [j], which would in its turn trigger the palatalization of the preceding /l(l)/. The final vocoid, having lost its time unit, would share one with the second segment of the preceding geminate, as in (50).

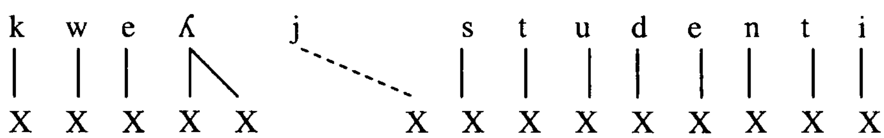
(50) Prosodic Reduction Palatilization



(modified from Burzio 1987:7)

Since forms like *begli* [bɛʎʎj] and *quegli* [kweʎʎj], lacking a final vowel, would not be well-formed, they only surface before a vowel - which can function as a peak for the ill-formed syllable – and when the following word undergoes pseudo-epenthesis. In the latter case, the final vocoid can function as the epenthetic vowel, associating with the empty time slot and reacquiring syllabic status. The interaction between 'prosodic reduction' and 'pseudo-epenthesis' is here perfectly parallel to that of the masculine singular, as can be seen in (51), parallel to (47) above.

(51) a. PROSODIC REDUCTION b. PSEUDO-EPENTHESIS



As for the forms *bei, quei, i,* Burzio (1987:9) advances the hypothesis that these surface when palatalization precedes, and therefore bleeds, gemination. The final [i] would in fact be the result of the palatalization of a single [l], while palatalization of [ll] would yield [ʎʎ]. This analysis, however, does not provide an explanation for this different ordering, nor does it account for the fact that the two opposite orderings apply in precisely complementary phonological environments. Burzio's (1987) account only explains why the forms *begli/ quegli/ gli* cannot appear before a consonant-initial word which has not undergone pseudo-epenthesis, but not why the forms *bei/ quei/ i* should be the alternative.[23]

More importantly, Burzio (1987) does not make explicit use of the notion of syllabic structure, although this is clearly at the basis of the processes at hand. In the characterization of both the reduction and the pseudo-epenthesis processes, he refers to time units rather than to subsyllabic nodes. By making reference to syllable structure, however, one can attain a more straightforward characterization of the process.

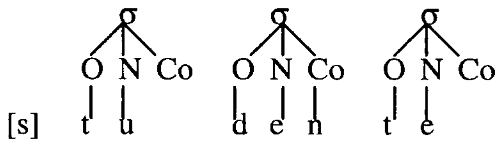
5.2.2.3.2 A directional approach

Pseudo-epenthesis, like most insertion processes, may be construed as intrinsically related to syllabification. A prosodic theory assuming directional parsing (Itô 1989, Noske 1993) actually provides a neat account of the phenomenon.

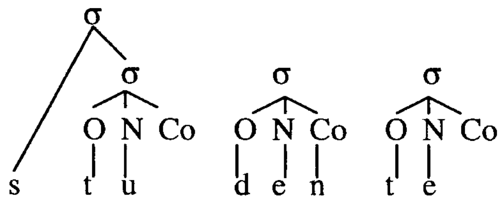
According to such an approach, in Italian syllabification applies from right to left, maximizing onsets (cf. 1.3). Therefore, in a word like *studente* the last syllable $[te]_\sigma$ is formed first, then the syllable $[den]_\sigma$ and finally the second and third segments are joined into $[tu]_\sigma$ (52a). Since /s/C is not a possible onset, the initial /s/ cannot be parsed into the latter syllable. If the word appears in isolation or is preceded by a consonant-final lexical item, the /s/ cannot be assigned to a syllable coda either: the degenerate syllable option is ruled out because there is no epenthetic vowel that could fill in its nucleus. We may suppose that at this point, for a parametrical choice of the language, the consonant is not deleted, but that it is incorporated into the following syllable (52b). As will be shown in 5.3, an analogous device was proposed by Mc Carthy (1979a) in the treatment of Arabic superheavy syllables. If, on the other hand, the cluster follows a vowel-final word such as *nessuno* (52c) or *caro* 'dear' (52d), the initial [s] is syllabified into the empty coda of the preceding syllable.

[23] Also see objections raised in 4.2.

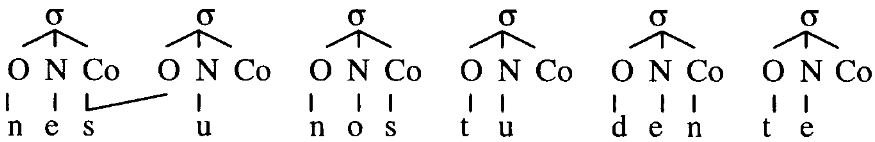
(52) (a) *Initial syllabification*



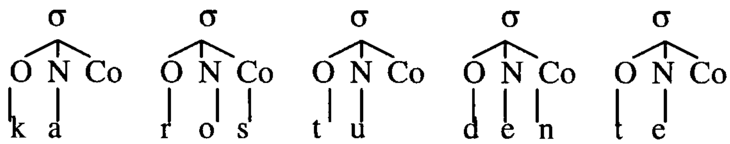
(b) */s/ incorporation*



(c) *nessuno studente* 'no student (m.)'



(d) *caro studente* 'dear student (m.)'



One of the advantages of this model is that it avoids postulating a process of pseudo-epenthesis, because the inclusion of the stray consonant into the coda of the preceding syllable falls out naturally from the syllabification process itself. Specifier Vowel Deletion (see 4.2), however, still has to be characterized as an independent phenomenon, since it cannot be directly derived from syllable structure constraints.

Clearly, these considerations apply to all words beginning by one of the phonotactically disallowed consonant clusters (i.e. /s/+C and plosive+C), or by one of the consonants that are always geminate in SI (i.e. /j/, /ɲ/, /ʎ/, /ʃ/, /ts/ and /dz/). Once again, it must be stressed that we are only talking about the latter variety, while other Italian accents may, and most often do, react in different ways. For instance, the 'learned' cluster /pn/ in the word *pneumatico* 'tyre', when preceded by the definite article, will be resolved as *lo pneumatico* in SI but as *il p[e]neumatico* or *il p[ə]neumatico* in most colloquial varieties.

In a directional approach, the preference for preservation of the /s/C and plosive + consonant clusters in isolation over other strategies (for instance, deletion of one of the consonants) must be stipulated. In OT terms, on the other hand, this preference can be expressed quite naturally, by postulating that, in SI, faithfulness constraints override the ban on certain onset clusters. This latter ban is expressed in terms of SON-SEQ in the case of /s/C (cf. 4.2), of SON-DIS in the case of plosive + consonant clusters.[24] For instance, tableau (53) shows that the faithfulness constraints MAX-C-IO, militating against CD, and DEP-V-IO, countering VI, must outrank SON-SEQ.

(53)

Input: /studɛnte/	MAX-C-IO	DEP-V-IO	SON-SEQ
a. ☞ [stu.dɛn.te]			*
b. [tu.dɛn.te]	*!		
c. [su.tu.dɛn.te]		*!	
d. [es.tu.dɛn.te]		*!	

When a word like *studente* 'student' is preceded by a vowel-final item such as *nessuno* 'no', the sonority-violating cluster is broken up without any need for contravening DEP-V-IO, since the pseudo-epenthetic vowel is already present in the input. As shown in tableau (54), all candidates exhibiting truncation are suboptimal, because they violate either MAX-C-IO, or SON-SEQ, or *COMPLEX^{COD}.[25]

(54)

Input: /nessuno studɛnte/	MAX-C-IO	DEP-V-IO	SON-SEQ	*COMPLEX^{COD}	ALIGN-L (PW, σ)
a. ☞ [nes.su:.nos.tu.dɛn.te]					*
b. [nes.sun.stu.dɛn.te]			*!		
c. [nes.suns.tu.dɛn.te]				*!	
d. [nes.sun.tu.dɛn.te]	*!				

The faithful optimal candidate, however, violates ALIGN-L (PW, σ), which must therefore be outranked by the other constraints (see 4.2 for discussion).

[24] The fact that the geminate consonants are always simplified in post-pausal or post-consonantal position indicates that *[...C_1C_1 ...]$_\sigma$, contrary to SON-SEQ and SON-DIS, is inviolate in Italian (see 4.2).

[25] Clearly, all of these candidates will also violate MAX-V-IO, but this constraint is, apparently, ranked rather low in Italian (cf. 4.2).

None of these constraints is relevant to phrases such as *nessun cane* 'no dog', in which the noun specifier precedes a word beginning with a simple consonant. Therefore, as was explained in 3.2 and 4.2, EVEN-TROCHEE is free to trigger truncation, by dominating MAX-V-IO.

5.3 Insertions in Arabic

This section deals with vocoid insertion phenomena in the Arabic dialects, with special reference to AA. In subsection 5.3.1, the process of Glide Insertion (GI), which is rather uncommon in varieties like AA, is analyzed. The following subsection deals instead with vowel epenthesis. It first intruduces the environment in which the phenomenon takes place, clarifying its connection with syllable structure. Interdialectal differences in the quality of the epenthetic vowel and especially in the epentesis sites are then analyzed. The last two subsections deal with the problematic interaction of epenthesis with syncope and stress assignment. In all cases, the various DT approaches are compared to an OT account of the phenomena.

5.3.1 Glide Insertion

As was briefly mentioned in 2.3, GI does not commonly apply across PW-boundaries in varieties such as AA. When a PW-final high vocoid clashes with a PW-initial vowel, either GF takes place, or ʔ-insertion does. That is, the insertion of the default contoid [ʔ] is generally preferred over the copying of the high vocoid into the following empty onset. Examples from AA are reported in (55).

(55) *zoojti [ʔ]urduníyya ~ zoojt[j] urduníyya* 'my wife is from Jordan'

inti [ʔ] abqaríyya ~ int[j] abqaríyya 'you (f.sg.) are a genius'

biddi [ʔ] ashúuf ~ bidd[j] ashúuf 'I want to see'

haada-l naadi [ʔ] árxaS 'this club (is) cheaper'[26]

The only clear cases of GI apply inside the PW. For instance, as was pointed out in 2.3 and 3.3, when a level 2 suffix like *–a* is added to the stem of nisba adjectives to form the feminine, the stem-final /I:/ is lengthened by one timing slot to fill in the

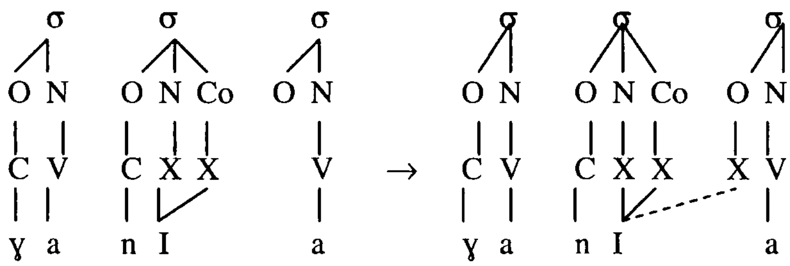
[26] The GF option is not available in this case because the second clashing vowel is stressed (see 2.3).

empty onset of the following syllable. The masculine and feminine of a few of these adjectives are reported in (56).[27]

(56)	ghan[i]	~	ghan[ijj]a	'rich'
	9aad[i]	~	9aad[ijj]a	'ordinary'
	bunn[i]	~	bunn[ijj]a	'brown'
	suur[i]	~	suur[ijj]a	'Syrian'
	ajnab[i]	~	ajnab[ijj]a	'foreign'

In DT, the process can be represented as in (57). Its domain of application seems to be the PW.

(57) /ɣanI:/ + /a:/ → [ɣanij.ja] 'rich (f. sg.)'



In an OT framework, the preference for GI over ʔ-insertion within the PW can be attributed to the ranking DEP-C-IO » DEP-X-IO. The dominant constraint states in fact that a consonant in the output must have a correspondent in the input, thus militating against CI, of which ʔ-insertion is a particular case. The insertion of a glide incurs a less serious violation of the lower-ranked DEP-X-IO, and is therefore preferred (see tableau (58)). Clearly, NOHIATUS must also outrank this constraint, in order to eliminate the faithful candidate c.[28]

[27] It must be borne in mind that, in the masculine forms, the /I:/ has undergone Final Vowel Shortening (3.3).
[28] For an evaluation of other possible candidates – as for example the one with GF – see 2.3.

(58)

Input: /ɣanI: + a/	NOHIATUS	DEP-C-IO	DEP-X-IO
a. ☞[ɣa.nij.ja]			*
b. [ɣa.ni:.ʔa]		*!	
c. [ɣa.ni:.a]	*!		

However, this ranking does not yield the correct result for the clashes at PW-boundaries, if some higher ranked constraint does not intervene to rule out GI. This cannot be a syllable or foot structure constraint, because the candidate with GI has exactly the same syllable and foot structure as that with ʔ-insertion (cf. *(naa.di)(jar.xaS)* and *(naa.di)(ʔar.xaS)* 'club (is) cheaper'). The difference between the two candidates lies instead in their segmental make-up. In order to express the notion that this segmental difference becomes relevant at PW-boundaries, the alignment constraint ALIGN-L (PW, seg) can be invoked. Such a constraint demands that the left edge of a PW coincide with the left edge of a segment. Since a PW-final vocoid that is lengthened to fill in the onset of the next syllable clearly straddles a PW-boundary, GI violates ALIGN-L (PW, seg). If this is ranked above DEP-C-IO, the candidate with ʔ-insertion, which does not violate ALIGN-L (PW, seg), will be selected. The evaluation of the candidates for *naadi [ʔ] árxaS* 'club (is) cheaper' is reported in (59).

(59)

Input: /na:di arxaS/	NOHIATUS	ALIGN-L (PW, seg)	DEP-C-IO	DEP-X-IO
a. $[na:.di]_{PW}$ $[jar.xaS]_{PW}$		*!		*
b. ☞$[na:.di]_{PW}$ $[ʔar.xaS]_{PW}$			*	
c. $[na:.di]_{PW}$ $[ar.xaS]_{PW}$	*!			

5.3.2 Vowel Insertion

In Arabic, VI often applies in order to break up phonotactically disallowed consonant clusters. The alternative strategy, i.e. deletion of one of the consonants, is generally avoided. This tendency may well be due to the fact that, in this language, consonants make up the root melody; their preservation thus acquires particular importance for the basic meaning of lexemes not to be impaired.

5.3.2.1 Positions of VI

Vowel epenthesis in Arabic may take place in four positions:

1. Post-pausally;
2. Word-medially;
3. Pre-pausally;
4. Between words.

5.3.2.1.1 Post-pausal insertion

In post-pausal position, epenthesis may apply to break up domain-initial consonant clusters by inserting a vowel before or after the leftmost consonant. Such clusters derive mainly from three sources: either they are part of certain verbal forms like the imperatives (e.g. /ktub/ 'write!'), or they result from morpheme concatenation (e.g. /l-kita:b/ 'the book'), or they are a consequence of syncope (see 4.3). Such a source often determines the frequency with which the clusters undergo VI: while the process is obligatory in the case of the imperatives (see AA *[u]ktub!* 'write!', *[i]Hki!* 'talk!'),[29] it is rather uncommon when the cluster is derived by means of VD. Moreover, the type and number of allowed initial clusters varies from dialect to dialect, with the Levantine varieties epenthesizing considerably more than the Maghrebian ones (which correspondingly delete more, see 4.3).

In AA, though branching onsets are never allowed in domain-medial position, two-consonant clusters are commonly found post-pausally (e.g. *klaab* 'dogs', *kbiir* 'big', *Hmaar* 'donkey'). Epenthesis, however, may occasionally take place, especially if the sequence is particularly difficult to pronounce, as in *qmaash* ~ *[i]qmaash* 'fabric' or *[i]tkallam[i]t* '(I) spoke'.

As was mentioned in 4.3, domain-initial consonant clusters which derive from syncope often seem to preserve a residue of the deleted vowel. This may be the reason why they rarely trigger epenthesis. In accordance with these considerations, Watson

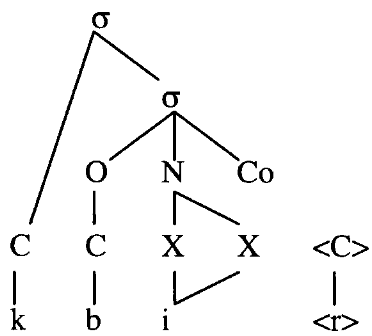
[29] This seems to be an altogether different process which, due to its obligatoriness and to its morphological and lexical conditioning, probably applies in the lexicon.

(1999a:518-21, 2000:22-3, 2002:76) assumes that the mora of the deleted vowel may be preserved and assigned to the initial consonant, as in (60).

(60) *kbiir* 'big'

Alternatively, such a consonant may be seen as an extrametrical segment, later adjoined to the following syllable by incorporation (cf. McCarthy 1979a, Selkirk 1981b). The representation of *kbiir* 'big' in (61) parallels that of the Italian example in 5.2.2.3.2.

(61)



5.3.2.1.2 Word-medial epenthesis

Epenthesis is instead very common word-medially, where it generally applies to break up three- and four-consonant clusters. In AA and in most of the Mashreq dialects, this type of sequences is generally given by the combination of two or more morphemes (e.g. AA *katab* + *t* + *l* + *ha* > *katabtilha* '(I) wrote to her'). While in the Maghreb such clusters may be preserved (cf. Cantineau 1960:114), in dialects like AA they systematically trigger epenthesis, as in (62).

(62) /katab/ + /t/ + /ha/ → [katabitha] '(I) wrote it (f.)'

/kalb/ + /na/ → [kalibna] 'our dog'

/bank/ + /hum/ → [banikhum] 'their bank'

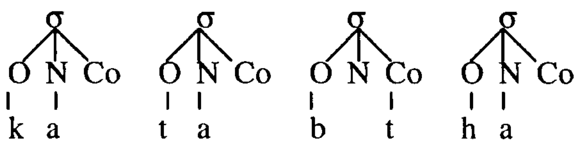
/ʔul/ + /t/ + /l/ + /ha/ → [ʔultilha] '(I) said to her'

/katab/ + /t/ + /l/ + /kum/ → [katabtilkum] '(I) wrote to you (pl.)'

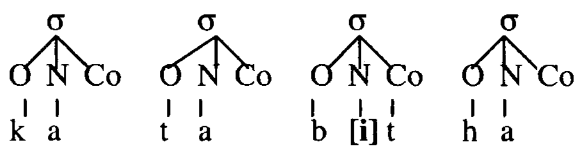
Since in AA both complex onsets and complex codas are completely banned (at least in word-medial position), a CVCCCV or CVCCCCV sequence cannot be parsed into well-formed syllables without the insertion of an extra nuclear element. If degenerate syllables are assumed as the output of initial syllabification, the process may be represented as in (63).[30]

(63) /katab-t-ha/ > [katabitha] '(I) wrote it (f.)'

(a) *Initial syllabification:*



(b) *Insertion of a default vowel:*



In an OT perspective, the major trigger of word-medial epenthesis is *COMPLEX, disallowing complex onsets and codas. This constraint will clearly outrank DEP-V-IO, which militates against VI. This latter must also be dominated by MAX-C-IO, which thus rules out CD as an alternative strategy. The evaluation of the candidates for *katabitha* '(I) wrote it (f.)' is given in tableau (64).

[30] As was explained in 3.3, dialects which do not always implement Closed Syllable Shortening may adopt epenthesis as a device to eliminate not only VCC, but also VVC rhymes word-internally (i.e. when the final consonant cannot be extrametrical). Examples from Makkan Arabic (Abu Mansour 1987, 1992) are *xaalana* 'our maternal uncle', *kitaabaha* 'her book'. AA, however, does not epenthesize in these cases, and tends to preserve the CVVC syllable (e.g. *xaalna* 'our uncle' or *kitaabha* 'her book').

(64)

Input: /katab-t-ha/	*COMPLEX	MAX-C-IO	DEP-V-IO
a. [ka.tabt.ha]	*!		
b. [ka.tab.tha]	*!		
c. ☞ [ka.ta.bit.ha]			*
d. [ka.tab.ha]		*!	

In this tableau, both candidate a. and candidate b. are eliminated by *COMPLEX, since the former exhibits a branching coda, the latter a branching onset. Candidate d., with a deleted consonant, violates instead MAX-C-IO. Ranking between these two constraints cannot be established on the basis of the data at hand.

5.3.2.1.3 Pre-pausal epenthesis

Pre-pausal VI applies to break up a domain-final two- or three-consonant cluster. In the modern dialects, the phenomenon is very frequent, especially since word-final clusters have increased in number because of diachronic VD (Cantineau 1960:14).

This type of epenthesis, however, does not apply in all dialects with the same frequency. In varieties like Makkan[31] and Cairene Arabic, the cluster tends to be preserved (e.g. *katabt* '(I) wrote'), while in Sudanese, Iraqi, Syrian, Levantine, Palestinian, VI often applies (e.g. Sudanese *katabta*, Iraqi *kitabit* '(I) wrote') (Abu-Mansour 1992, Broselow 1992, Majdi and Winston 1994).

AA, on the other hand, generally epenthesizes only when the cluster is difficult to pronounce, i.e. when it does not respect sonority sequencing, as in the examples in (65).

(65) *'abil* 'before, ago'
xubiz 'bread'
jisir 'bridge'
baTan 'stomach'
'udun 'ear'
laHem 'meat'
baHer 'sea'
jabir 'algebra'
sifir 'book of the Torah'

[31] In Makkan Arabic, pre-pausal epenthesis applies only to a restricted group of items (mostly nouns), such as *is[i]m* 'name' or *shuk[u]r* 'thanks' (Abu-Mansour 1992:139).

'asir 'capture'

Coda sequences that decrease in sonority are instead generally preserved (66).

(66) *ba9d* 'after'
bard 'cold'
9ind 'at, to'
bank 'bank'
wasT 'centre'
gharb 'west'
law samaHt 'please'
Darf 'envelope'
kanz 'treasure'
shirk 'polytheism'
kalb 'dog'

Examples such as *biZZabt* 'exactly', *shaxS* 'person', *katabt* '(I) wrote' or *nafs* 'same' indicate that clusters formed by consonants of equal sonority also tend to be preserved.

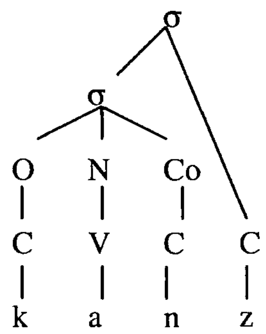
In spite of what was said above, a vowel is occasionally inserted to break up a cluster even when it does not violate sonority sequencing, as in *daris* 'lesson', *itkallamit* '(I) spoke', *miliH* 'salt'. Other words very rarely undergo epenthesis even if they exhibit sonority-violating clusters (e.g. *maSr* 'Egypt', *xamr* 'alcohol', *tamr* 'date (fruit)').

Once again, the process applies in order to avoid phonotactically disallowed coda sequences. However, while word-internally complex codas are altogether banned, in pre-pausal position sequences of two consonants are often allowed, especially if they do not violate sonority sequencing.[32] For those cases in which epenthesis does not apply, the final consonant can be construed as extrametrical. Just like in the case of

[32] Selkirk (1981b:220-1) points out that the existence of superheavy syllables pre-pausally is perfectly compatible with the phonetic tendency to lengthen syllables in the same position (see 3.3). While phonological processes are generally structure-preserving, at the limit of prosodic domains (in this case, of the PU) rules producing structures which are not present at the underlying level (in this case, superheavy syllables) may apply.

post-pausal clusters, such a segment is incorporated into the preceding syllable at a later stage, as in (67).

(67)



(adapted from McCarthy 1979a, Selkirk 1981b)

Selkirk (1981b) actually assumes that the final consonant, rather than being extrametrical, is assigned to a degenerate syllable by initial syllabification. The CVC.C__ (or CV.C__C) sequence would then be collapsed into a single heavy syllable CVCC, thus bleeding epenthesis. Abu-Mansour (1992:149-52), however, shows that a degenerate syllable approach encounters difficulties in dealing with stress assignment in polysyllabic forms. In words such as /katab-t/ > [ka'tabt] '(I/you m.sg.) wrote', epenthesis - which in this case fails to apply, bled by syllable incorporation - must be ordered before stress assignment. The reverse order would yield the wrong surface form ['ka.tabt], deriving from an intermediate |'ka.ta.bΔt|, where [bΔt] is the degenerate syllable. If, on the other hand, the final consonant is extrametrical, stress will be automatically assigned to the heavy syllable [tab], without any need for rule ordering.

In cases in which VI applies, however, phonotactic considerations must prevail over syllable structure considerations. Abu-Mansour (1992) argues that, while word-medial and post-pausal epenthesis are syllabically conditioned, pre-pausal epenthesis is segmentally conditioned in Makkan Arabic. This dialect, just like AA, only breaks up clusters which violate sonority sequencing. According to Abu-Mansour (ibid.), epenthesis applies to consonants which have already been assigned to well-formed syllables, since Makkan Arabic allows word-final two-consonant clusters. Therefore, pre-pausal epenthesis must take place at a later stage in the derivation than medial and post-pausal epenthesis.

In an OT perspective, the difference between those forms which undergo epenthesis and those which do not may be expressed in terms of the constraint SON-SEQ: where this dominates DEP-V-IO, sonority-violating clusters trigger epenthesis. The different behaviour of domain-final and domain-medial clusters is instead given by high ranking of *FINAL-C-μ. As was mentioned in 1.5, this constraint requires a domain-final consonant to be weightless, thus translating into OT terms the notion of extrametricality. In order for final consonants to be extrametrical, *FINAL-C-μ must outrank WEIGHT-BY-POSITION (WBP), which requires coda consonants to be moraic.

The tableau in (68), in which non-moraic (i.e. 'extrametrical') consonants are in angled brackets, reports the evaluation of the candidates for *xubiz* 'bread', whose sonority-violating cluster is broken up by epenthesis.

(68)

Input: /xubz/	*FINAL-C-μ	WBP	*COMPLEX	MAX-C-IO	SON-SEQ	DEP-V-IO
a. xubz	*!		*			
b. xub<z>		*			*!	
c. xu<b>		*		*!		
d. xub[i]z	*!					
e. ☞ xub[i]<z>		*				*

Candidates a. and d., having moraic final consonants, violate the highly ranked *FINAL-C-μ, and are thus ruled out. The remaining candidates, all flouting WBP, have an extra violation each: MAX-C-IO for candidate c., SON-SEQ for candidate b. and DEP-V-IO for candidate e. In order for the latter to be the optimal output, DEP-V-IO must be outranked by the other two constraints. Since, however, epenthesis is optional, free ranking must be postulated between SON-SEQ and DEP-V-IO.

As shown in tableau (69), when a domain-final cluster decreases in sonority, SON-SEQ becomes irrelevant. While *FINAL-C-μ still eliminates a. and d., the candidate with an extrametrical consonant but without deletion or epenthesis is selected.

(69)

Input: /kanz/	*FINAL-C-μ	WBP	*COMPLEX	MAX-C-IO	SON-SEQ	DEP-V-IO
a. kanz	*!		*			
b. ☞ kan<z>		*				
c. ka<n>		*		*!		
d. kan[i]z	*!					
e. kan[i]<z>		*				*!

In word-medial position, where *FINAL-C-μ does not exert its influence, three-consonant clusters are broken up even when they would not violate sonority sequencing. The tableau for *kanizna* 'our treasure' in (70) illustrates how the epenthetic candidate is selected. With *FINAL-C-μ out of the way, WBP rules out any candidate with an extrametrical consonant. Candidates a. and b., on the other hand, have a complex coda and a complex onset respectively, and thus violate high-ranked *COMPLEX. As in the other cases, MAX-C-IO rules out the candidate with CD.

(70)

Input: /kanz + na/	*FINAL-C-μ	WBP	*COMPLEX	MAX-C-IO	SON-SEQ	DEP-V-IO
a. kanz.na			*!			
b. kan.zna			*!			
c. kan<z>.na		*!				
d. kan.na				*!		
e. ☞ ka.n[i]z.na						*

5.3.2.1.4 Insertion between words

In many Arabic dialects, epenthesis also applies when a three- or four-consonant cluster is formed by word concatenation. In this case, the three (or four) consonants belong to two or more lexemes, as in the AA examples in (71).

(71) /ʒabal/ + /kbI:r/ → [ʒabalikbi:r] '(a) big mountain'

/bIddak/ + /kta:b/ → [biddakikta:b] 'you (m.sg.) want (a) book'

/bIddak/ + /l/ + /kta:b/ → [biddaklikta:b] 'you (m.sg.) want the book'

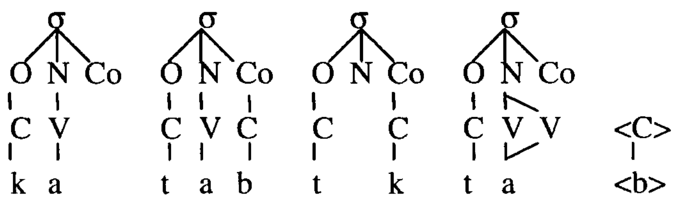
/katabt/ + /kta:b/ → [katabtikta:b] '(I) wrote (a) book'

/katabt/ + /l/ + /rIsa:la/ → [katabtirrisa:la] '(I) wrote the letter'

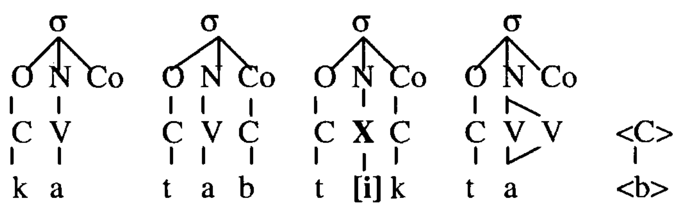
This is clearly a case of domain-medial VI, analogous to that breaking up word-internal clusters (see 5.3.2.1.2). Three- and four-consonant sequences cannot be exhaustively syllabified, since at least one complex onset or one complex coda would be formed. The extra consonants are therefore assigned to a degenerate syllable (or are left stray), and an epenthetic vowel is inserted. The process is exemplified in (72).

(72) /katabt/ + /kta:b/ —> [katabt**i**kta:b] '(I) wrote (a) book'

(a) *Initial syllabification:*



(b) *Insertion of a default vowel:*



However, contrary to word-internal epenthesis, in AA this type of VI does not apply every time a cluster is formed across a word-boundary. As the examples in (73) illustrate, when a word-final cluster that does not violate sonority is combined with a word-initial consonant to form a three-consonant sequence, epenthesis does not generally apply.

(73) /katabt/ + /rIsa:la/ —> [kata**btr**isa:la] '(I) wrote (a) letter'

/ʃUft/ + /kalb/ —> [ʃ**uftk**ɛlb] '(I) saw (a) dog'

/bInt/ + /ʒamI:la/ —> [bi**ntʒ**ami:la] '(a) beautiful girl'

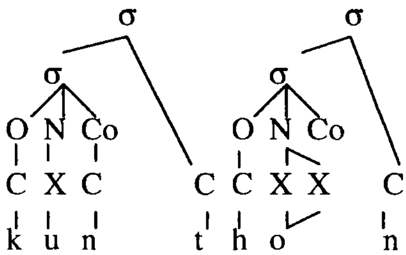
/katabt/ + /dars/ —> [kata**btd**aris] '(I) wrote (a) lesson'

/kUnt/ + /hU:n/ —> [ku**ntho**:n] '(I) was here'

It is possible that, in AA, two distinct VI processes take place: one inside the CG and one across CG boundaries, in a larger domain which may be identified with the PPh, as in Cairene (Majdi and Winston 1994:187) or in San9ani (Watson 1999a:509). The first process would be obligatory, the second one optional. However, VI does apply obligatorily when the second word begins with a cluster.

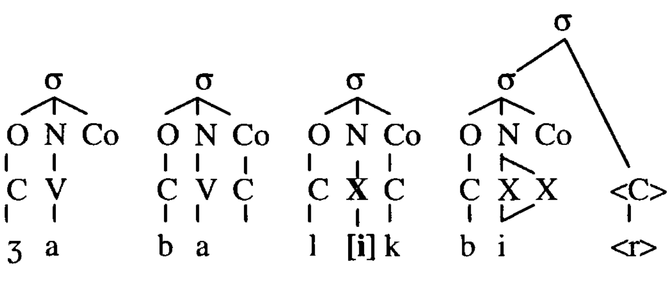
A more plausible account postulates that final extrametricality is a CG phenomenon, so that all CG-final consonants undergo incorporation and thus do not require epenthesis, as shown in (74).

(74) /kUnt/ + /hU:n/ → [ku**ntho**:n] '(I) was here'



Initial extrametricality, on the other hand, must refer to the PPh, since a CG-initial cluster triggers epenthesis when preceded by a consonant-final word[33] within the same PPh (cf. (75)).

(75) /ʒabal/ + /kbI:r/ → [ʒaba**li**kbi:r] '(a) big mountain'



In this respect, AA behaves differently from other Mashreq dialects such as Cairene, Iraqi or Sudanese, in which CG-final clusters do trigger epenthesis (cf. Cairene *katabt[i] gawaab* 'I wrote a letter', Sudanese *katabt[a] kitaab* 'I wrote a book', Iraqi

[33] A vowel-final word would instead trigger resyllabification of the first consonant of the cluster into the coda of its last syllable, thus bleeding epenthesis (e.g. *Tawla kbiira* [Tau̯.lak.bi:.ra] '(a) big table').

kitab[i]t maktuub 'I wrote a letter', from Broselow 1992:21). In these dialects, final extrametricality must be a PPh phenomenon, just like initial extrametricality.

5.3.2.2 Quality and location of the epenthetic vowel

5.3.2.2.1 Quality

The quality of the epenthetic vowel varies from dialect to dialect. For instance, it is [a] in San9ani, Makkan and Sudanese, [i] in Iraqi and Cairene, [ə] in Syrian. In AA, the epenthetic vowel is generally an [i] or [ɪ], although when it is extra-short it tends to resemble a *schwa*. After pharyngeals and emphatics, moreover, it is closer to [ε] (cf. (76b)).

(76) (a) jab[ɪ]r 'algebra'
xub[ɪ]z 'bread'
katabt[i]lha '(I) wrote to her'
ma kaan[i]sh '(he) was not'
(b) laH[ε]m 'meat'
taH[ε]t 'under'
baT[ε]n 'stomach'

In some cases, the epenthetic vowel may also be a (partial or complete) copy of a neighbouring vowel. In AA, as in Cairene (Watson 2000:12), the epenthetic vowel is [u] (or [o]) rather than [i] in the environment of another [u] (e.g. *ktub* > *[o]ktub* 'write (m. sg.)!', *'ud[u]n* 'ear').

In an OT framework, featural markedness constraints account for the choice of the epenthetic vowel. For instance, in those dialects in which the default vowel is [i], the context-free constraints *[+low] and *[+round] are highly ranked; where it is [a], the constraints *[+front] and *[+round] will dominate *[+low]. Where all these constraints, and also *[+high] and *[+back], outrank their negative counterparts, the epenthetic vowel will be [ə]. In cases like AA *'ud[u]n* 'ear', the constraint HARMONY(round), requiring a vowel to harmonize for roundness with neighbouring vowels, dominates *[+round] and *[+back]. Since epenthetic segments lack input correspondents, the action of IO-faithfulness constraints does not affect them.

5.3.2.2.2 Epenthesis sites

The choice of the insertion site, when more than one is possible, also varies across the Arabic dialects. In particular, a CCC sequence surfaces as CCVC in varieties such as Cairene, Sudanese and Makkan, as CVCC in varieties such as Iraqi, Syrian and Palestinian (Broselow 1992:22). The examples below, which have become classics in Arabic phonology, are from Cairene (77) and Iraqi Arabic (78). They show that, while the insertion site varies in the case of triconsonantal clusters ((77a) and (78a)), quadriconsonantal clusters are broken up between the second and third consonant in both dialects ((77b) and (78b)).

(77) *Cairene Epenthesis:* 0 > i / CC___C; 0 > i / CC___CC

(a) /ʔul-t-l-u/ → [ʔul.**ti**.lu] 'I said to him'

(b) /ʔul-t-l-ha/ → [ʔul.**til**.ha] 'I said to her'

(78) *Iraqi Epenthesis:* 0 > i / C___CC; 0 > i / CC___CC

(a) /gil-t-l-a/ → [gi.**lit**.la] 'I said to him'

(b) /gil-t-l-ha/ → [gil.**til**.ha] 'I said to her'

(from Itô 1989:242)

In domain-final position, a CC cluster is realized as CCV in Cairene-type[34] dialects, as CVC in Iraqi-type dialects (cf. Sudanese *katabta* vs. Iraqi *kitabit* '(I) wrote').
AA, like other Levantine varieties, appears to belong to this latter group. As is shown in (79), in this dialect a vowel is inserted between the first and second consonant of a CCC cluster.

(79) /kalb-na/ > [kal**i**bna] 'our dog' (cf. Cairene [kalb**i**na])

/bank-na/ > [ban**i**kna] 'our bank' (cf. Sudanese [bank**a**na])

/katab-t-ha/ > [katab**i**tha] 'I wrote it (f.)' (cf. Makkan and Sudanese [katabt**a**ha])

[34] As was mentioned in 5.3.2.1.3, however, Cairene Arabic tends to preserve the cluster in pre-pausal position.

/qabl/ + /ma:/ > [ʔabilma] 'before' (cf. San9ani [gablama:])

/ʃUf-t-nI/ > [ʃufitni] 'you (m.sg.) saw me' (cf. Makkan [ʃUftanI])

/ʕUmr-ha/ > [ʕumurha] 'her age' (cf. Makkan [ʕumraha])

The examples in (80) illustrate that in AA, when a domain-final cluster is not preserved, a vowel is inserted between its two consonants.

(80) /Ism/ > [ʔisim] 'name'

/ʃUkr/ > [ʃukur] 'thanks'

/TIfl/ > [Tɪfɪl] 'child'

/dars/ > [daris] 'lesson'

/ʔIbn/ > [ʔibɪn] 'son'

As for four-consonant clusters, they are invariably broken up between the second and the third segment, as in all other dialects (81).

(80) /katab-t-l-risa:la/ > [katabtirrisa:la] 'I wrote the letter'[35]

/ʔul-t-l-ha/ > [ʔultilha] 'I said to her'

/katab-t-l-ha/ > [katabtilha] 'I wrote to her'

/katab-t-l-na/ > [katabtilna] 'you (m.sg.) wrote to us'

/katab-t-l-kum/ > [katabtilkum] '(I) wrote to you (pl.)'

/bidd-ak-l-kta:b/ > [biddaklikta:b] 'you (m.sg.) want the book'

/katab-t kta:b/ > [katabtikta:b] 'I wrote a book'

/ʒa:b-l-ha/ > [ʒa:bilha] 'he brought to her'[36]

Triconsonantal clusters in pre-pausal position behave like quadriconsonantal clusters, in that they have two unsyllabifiable consonants. It is between these two consonants (the second and third) that the vowel is inserted, in all dialects. Examples from AA are reported in (81).

[35] Some phonologists (e.g. Selkirk 1981b) consider the [i] in [irrisa:la] 'the letter' as underlying.
[36] VVCCC sequences behave like four-consonant clusters because they exhibit two, rather than one, unsyllabifiable consonants, unless Closed Syllable Shortening applies (cf. 3.3).

(81) /ma/ + /kun-t-ʃ/ > [makuntiʃ] '(I/you m.sg.) was/were not'

/ma + /katab-t-ʃ/ > [makatabtiʃ] '(I/you m.sg.) did not write'

/ma/ + /ʔul-t-ʃ/ > [maʔultiʃ] '(I/you m.sg.) did not say'

In a particular group of words, AA appears to pattern like Cairene, rather than like Iraqi Arabic (see (82)).

(82) /ʔul-t-l-u/ > [ʔultillu] 'I said to him'

/ʔul-t-l-ak/ > [ʔultillak] 'I said to you (m.sg.)'

/katab-t-l-u/ > [katabtillu] 'I wrote to him'

/katab-t-l-ak/ > [katabtillak] 'I wrote to you (m.sg.)'

/difaʕ-t-l-ak/ > [difaʕtillak] 'I paid you (m.sg.)'

The above examples, however, belong to a well defined group of items, i.e. to those containing the prepositional clitic /-l-/. Contrary to what happens in Cairene (cf. *'ultilu* '(I) said to him'), in AA the consonant of such a clitic is doubled. The relevant sequences thus behave like four-consonant clusters, so that a vowel is inserted between the second and third consonant.
In a 'Skeletal Rule' approach, the epenthesis site is stated in the rule. AA epenthesis is thus formulated as in (83a), or in (83b).

(83) (a) 0 → i / C___C #
C

(b) 0 → [i] / ___C'; 0 → [i] / C' ___C' (C' = stray consonant)

As was mentioned in 5.1, in a framework assuming that epenthesis stems directly from initial syllabification, two main models have been proposed, both based on the assumption of degenerate syllables.[37]

[37] Abu-Mansour (1992), making reference specifically to epenthesis in Makkan Arabic, argues against an account postulating degenerate syllables and in favour of allowing unsyllabified (i.e. stray) consonants as the output of initial syllabification. He (ibid.:147) points out that a degenerate syllable

One such model, introduced by Selkirk (1981b) and developed by Broselow (1992), may be called the 'Onset/Rime' approach (Broselow 1992), because it postulates that the position of the epenthetic site depends on whether a stray consonant is syllabified in an onset (Cairene type) or coda (Iraqi type) position. In this perspective, AA is a 'rime-dialect'. Syllabification of a word like AA *kalibna* 'our dog' would thus take place as follows:

1. Initial syllabification applies, and the consonant [b], which cannot be syllabified without creating a complex onset or coda, is left stray.
2. The stray consonant is assigned a degenerate syllable $[l\_\mathbf{b}]_\sigma$, of which it occupies the coda.
3. The nucleus of the degenerate syllable is filled with the epenthetic vowel.

In the case of quadriconsonantal clusters, however, this approach would yield the wrong results. Since two consonants are left stray, they should both be syllabified as onsets or both as codas of degenerate syllables, calling for the insertion of two vowels instead of one. In order to account for these cases, Selkirk (1981b:218) postulates that, as a general principle, the number of dummy elements (i.e. of epenthetic positions) in a string must be minimized: under this assumption, no other analysis is available but the one that inserts a single vowel between the two stray consonants. Itô (1989), however, points out that this device is too powerful, in that it requires 'global computational power'.

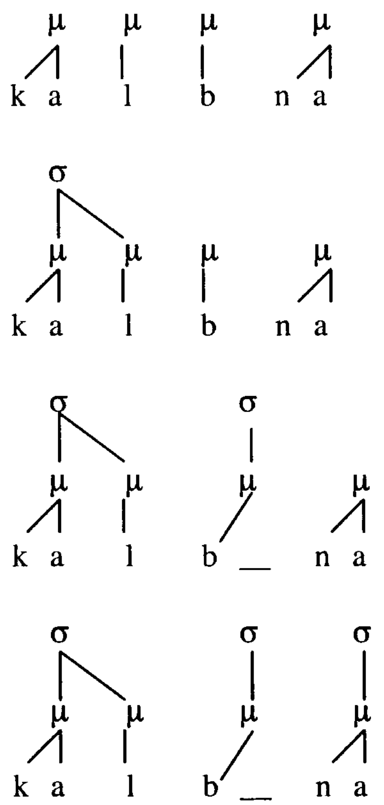
Broselow (1992) puts forward essentially the same claims as Selkirk (1981b), only within a moraic framework. She argues, more specifically, that the onset/rime distinction is due to a parametrical choice by which the rime dialects would allow a nuclear mora to dominate a consonant, at least at the underlying level, while the onset dialects would not (see 4.3). In the latter, therefore, in order to be saved from Stray Erasure, an extra consonant has to be incorporated into a degenerate syllable and linked directly to the syllable node, so that it becomes an onset consonant. In the rime dialects, on the other hand, at least at some level of representation a consonant may constitute a syllable. In varieties like Iraqi and AA, however, consonants cannot be syllabic in surface representations, so that 'a consonantal nucleus is interpreted as

analysis obtains the wrong results in cases in which the first of two stray consonants is preceded by a long vowel; a word like Makkan /sa:b-l-ha/ 'he poured for her' would surface as *[sa:.bal.ha] rather than the correct [sab.la.ha]. In Abu-Mansour's model, epenthesis creates a new syllable in which the stray consonant is syllabified, and whose nuclear slot will be later filled in by a default vowel.

a vowel-consonant sequence by rules of phonetic implementation' (Broselow 1992:31).

As was mentioned in 5.1, the second prosodic model that has been adopted to account for the different epenthetic sites in the Arabic dialects is based on directional syllabification (see especially Itô 1989). Assuming a moraic approach, Itô (1989) argues that Cairene-type dialects have left-to-right syllabification, Iraqi-type dialects right-to-left syllabification. A word like Cairene *kalbina* 'our dog' will therefore be syllabified as in (84).

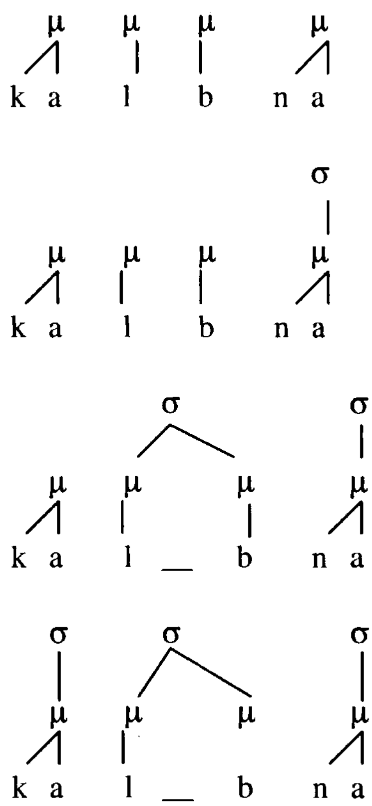
(84) (a) (b) (c) (d)



To begin with, moraic structure is built, assigning a mora to each CV pair and one to each consonant which does not have a vowel to its right (84a). The first two morae are then joined to form a syllable, since they conform to sonority sequencing and do not contain more than one nuclear element (84b). The second two morae, however, cannot be syllabified together, because this would result in a complex onset. The third mora is thus syllabified by itself but, since it comprises only a consonant, a dummy nucleus has to be inserted, which will be later filled in with an epenthetic vowel (84c). The last mora makes up a syllable by itself (84d).

The corresponding form in AA, *kalibna* 'our dog', is obtained by means of right-to-left syllabification, as in (85).

(85) (a) (b) (c) (d)



First, moraic structure is assigned (85a). The last two moras are then taken into consideration, but they cannot be joined into a syllable because of the ban against complex onsets. The last mora thus forms a syllable on its own (85b). The second and third mora can instead be syllabified together, provided an empty nucleus – later filled in by epenthesis - is inserted between them (85c). Finally, the initial mora is syllabified (85d).

Itô (1989:246-7) then shows that, with this method, quadriconsonantal clusters are always broken in the middle, whatever the direction of syllabification, since Maximality dictates the formation of a heavy syllable. She then points out that, although maximizing the size of the single syllables and minimizing the total number of syllables may seem to be equivalent, Maximization is to be preferred, in that it operates locally and does not call for a comparison of the candidate surface outputs of the entire string. More importantly, the position of the extra consonant within the degenerate syllable does not need to be stipulated, but falls out naturally from directional parsing. Moreover, a directionality analysis predicts the correct insertion sites for strings of five intervocalic consonants,[38] while an Onset/Rime analysis does not, even with the help of number minimization (Itô 1989:248).

As was mentioned in 5.1, within an OT framework the location of epenthesis sites seems to be best accounted for in terms of alignment constraints. In 'left-to-right', or 'onset', dialects, the constraint ALIGN-R(σ, PW) must be highly ranked. As the tableau for Cairene *kalbina* 'our dog' in (86) shows, if this constraint is ranked above ALIGN-L(σ, PW) the candidate in which the cluster is broken up between the second and third consonant (i.e. c.) is selected. This latter, in fact, has fewer violations for ALIGN-R(σ, PW) than candidate d., exhibiting epenthesis between the first and the second consonant.

(86)

Input: /kalb-na/	*COMPLEX	ALIGN-R (σ, PW)	ALIGN-L (σ, PW)
a. [kal.bna]	*!	μ	μμ
b. [kalb.na]	*!	μ	μμμ
c. ☞ [kal.**bi**.na]		μ μμ	μμ μμμ
d. [ka.**li**b.na]		μ μμμ!	μ μμμ

Candidate d., on the other hand, is selected in Iraqi-type dialects like AA, in which ALIGN-L(σ, PW) outranks ALIGN-R(σ, PW). The tableau for AA *kalibna* 'our dog' is reported in (87).

(87)

Input: /kalb-na/	*COMPLEX	ALIGN-L (σ, PW)	ALIGN-R (σ, PW)
a. [kal.bna]	*!	μμ	μ
b. [kalb.na]	*!	μμμ	μ
c. [kal.**bi**.na]		μμ μμμ!	μ μμ
d. ☞ [ka.**li**b.na]		μ μμμ	μ μμμ

5.3.2.3 Vowel Insertion and Vowel Deletion

VI and VD, being mirror-image processes, may be seen as two sides of the same coin, i.e. as one and the same vowel ~ 0 alternation (see 4.1 and 5.1). This is the point of view that Itô (1989) appears to assume although, as was mentioned in 4.3,

[38] These do not appear in Arabic but do in other languages, such as Temiar.

not distinguishing between underlying vowels which are deleted and epenthetic vowels which are absent from the UR may become problematic in specific cases. Leaving this issue for further research, it is however worth pointing out how the two processes often interact in the Arabic dialects.

5.3.2.3.1 Interaction of the two processes

It has been argued that epenthesis often feeds syncope by creating two consecutive light syllables. The examples in (88) are from Cairene Arabic, but they have identical correspondents in AA.

(88) /bInt/ + /kIbI:ra/ > [binti**kbi:ra] '(a) big girl'
/HUbb/ + /kItI:r/ > [Hubb**i**kti:r] 'a lot of love'

(from Watson 2000:18)

This has often been interpreted as VI preceding VD in the derivation (89a), since, if VI followed VD, the environment for this latter to apply would not be met (89b).

(89) (a) UR: /bint kibi:ra/ '(a) big girl'
VI: bint**i** kibi:ra
VD: binti kbi:ra
PF: binti kbi:ra

(b) UR: /bint kibi:ra/
VD: NA
VI: bint**i** kibi:ra
PF: *binti kibi:ra

In dialects such as Palestinian (Kenstowicz 1981:21) and AA, however, VD often seems to supply the environment for VI, by originating consonant clusters. The examples in (90) are from AA.

(90) /Haʒar/ + /kIbI:r/ > [Haʒar**i**kbi:r] 'a big stone' (cf. Cairene [Hagarkibi:r])

/Sadda:m/ + /HUsaIn/ > [Sadda:**mi**Hse:n] 'Saddam Hussayn'

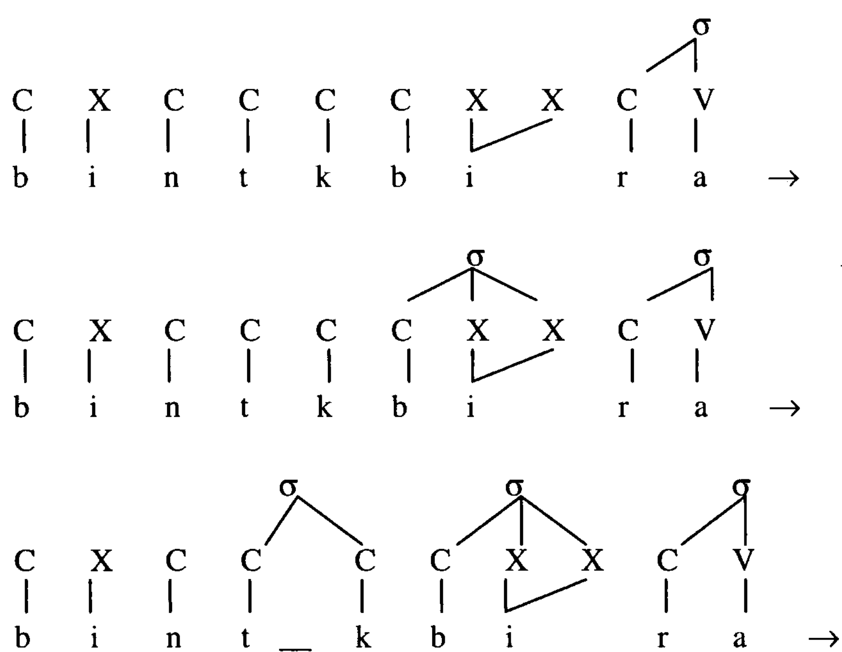
/ha:da/ + /l/ + /kIta:b/ > [ha:da**li**kta:b] 'this (is) the book'

Here, VI must be fed by VD, because a sequence like /Haʒarkibi:r/ does not exhibit any three-consonant clusters, and all of its segments can be syllabified without any need for epenthesis. The cluster [rkb] only originates with the application of syncope, which deletes the first /I/ of /kIbI:r/.

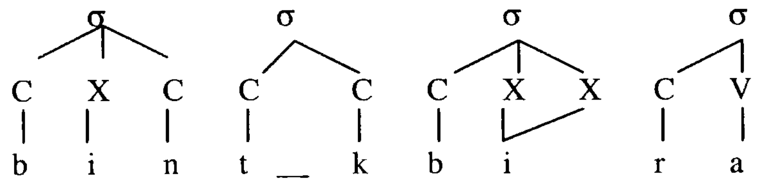
In a traditional, rule-based DT framework this ordering paradox may be explained by making VD and VI apply cyclically, so that VD applies first to /kIbI:r/ at the word level and then feeds VI at the phrase level. Alternatively, as was mentioned in 4.3, VD may have been lexicalized in items such as *kbiir* 'big' and *ktiir* 'a lot'.

A directional approach, however, can easily account for all of the above examples by postulating right-to-left syllabification in AA (see 5.3.2.2.2).[39] Under this analysis, a phrase like *bint ikbiira* '(a) big girl' is syllabified as in (91), where the correct result is obtained.

(91) ⇐ direction of syllabification



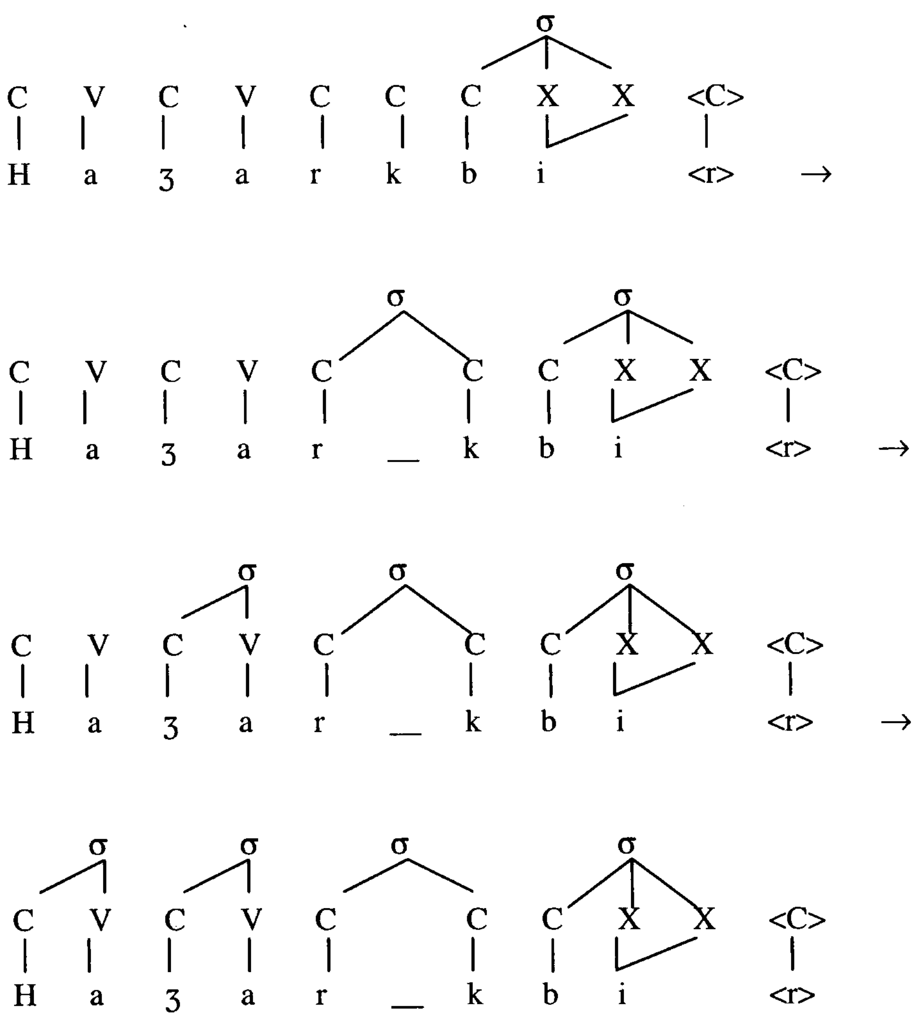
[39] As was explained in 4.3, Broselow (1992) accounts for similar cases by assuming that Iraqi-type dialects allow consonants to be syllabic at some level of representation.



Right-to-left syllabification also yields the correct epenthesis site for phrases such as *Hajar ikbiir* '(a) big stone' (92).

(92)

⇐ direction of syllabification



In an OT framework adopting Generalized Alignment, the same results follow straightforwardly from the assumption that, in dialects like AA, ALIGN-L (σ, PW) is highly ranked. More specifically, this constraint can account for the above choices if it dominates both MAX-V-IO and DEP-V-IO (see tableaux (93) and (94)). In tableau (93), since the faithful candidates have either a complex onset or a complex coda, the constraint *COMPLEX has also been taken into consideration. The faithful candidate in tableau (94), on the other hand, does not violate *COMPLEX, which has therefore been left out.

(93)

Input: /bint kIbI:ra/	*COMPLEX	ALIGN-L (σ, PW)	MAX-V-IO	DEP-V-IO
a. [bint.ki.bi:.ra]	*!	μμμ μμμμ μμμμμμ		
b. [bin.tki.bi:.ra]	*!	μμ μμμ μμμμμ		
c. [bi.nit.ki.bi:.ra]		μ μμμ μμμμ μμμμμ! μ		*
d. [bin.ti.ki.bi:.ra		μμ μμμ μμμμ μμμμ! μμ		*
e. ☞ [bin.tik.bi:.ra]		μμ μμμμ μμμμμμ	*	*

(94)

Input: /Haʒar kIbI:r/	ALIGN-L (σ, PW)	MAX-V-IO	DEP-V-IO
a. [Ha.ʒar.ki.bi:r]	μ μμμ μμμμ!		
b. ☞ [Ha.ʒa.rik.bi:r]	μ μμ μμμμ	*	*

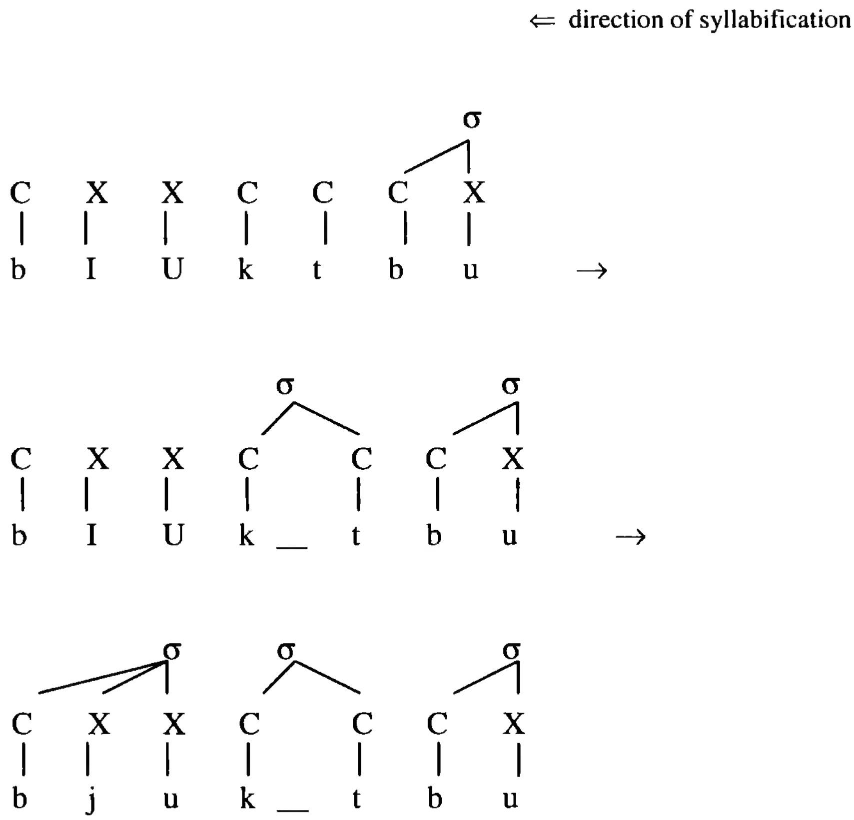
5.3.2.3.2 A problem with all these approaches

As was shown in 4.3, however, AA does not always pattern like Iraqi or Syrian Arabic. Words like /bIUktUbU/ '(they) write' are realized in AA as [bjuktubu], rather than the expected [bjuk**u**tbu] (cf. Syrian: [bjəkətbu]). This constitutes a problem for all the above-mentioned accounts.

In a directionality framework, the surface form [bjuktubu] can only be obtained if syllabification applies from left to right, as in Cairene.[40] If right-to-left syllabification is postulated, the wrong form *[bjuk**u**tbu] is obtained (see (95)).

[40] Hence, the conclusion reached in 4.3 that AA is a structure-preserving, Cairene-type dialect.

(95)



Likewise, Selkirk's (1981b) and Broselow's (1992) onset/rime approach would fail to account for the difference between AA *Hajarikbiir* '(a) big stone' and *byuktubu* '(they) write'. According to the first example, AA would be classified as a rime dialect, according to the second, as an onset dialect.

The same problem would affect an OT analysis based on Generalized Alignment. If ALIGN-L (σ, PW) dominates ALIGN-R (σ, PW), as was postulated to account for the AA examples above, the wrong candidate is selected (b. in tableau (96)).

(96)

Input: /bI-UktUb-U/	ALIGN-L (σ, PW)	ALIGN-R (σ, PW)	MAX-V-IO	DEP-V-IO
a. [bjuk.tu.bu]	μμ μμμ!	μ μμ		
b. ☞ [bju.kut.bu]	μ μμμ	μ μμμ	*	*

It is clear from tableau (96) that candidate a. would be selected if ALIGN-R (σ, PW) (or MAX-V-IO, or DEP-V-IO) outranked ALIGN-L (σ, PW).

Zawaydeh (1997) points out that similar counterexamples may be found in Cairene and Sudanese. Since Cairene behaves as a 'left-to-right' (or 'ALIGN-R') language as regards epenthesis, the expected output for a phrase like /ma-bi-ni-ʃu:f-hum-ʃ/ 'we don't see them' should be *[mabniʃufhumʃ]. The actual surface form is instead [mabinʃufhumʃ], which would indicate right-to-left syllabification or high ranking of ALIGN-L (σ, PW). The same can be said for some syncopated forms in Sudanese, otherwise a 'Cairene-type' dialect.

Since an in-depth study of this apparent paradox is outside the scope of this thesis, we can provisionally conclude, with Zawaydeh (1997), that the outputs for epenthesis and syncope may be independent. Clearly, this is a problem especially for output-oriented OT.

5.3.2.4 Interaction with stress assignment

As was mentioned in 5.1, epenthetic vowels are often unstressed, even if they occupy a position to which prominence should be assigned.

In the Palestinian Arabic forms in (97), for instance, the epenthetic vowel belongs to a penultimate heavy syllable, and should therefore be stressed. Prominence is instead assigned to the preceding light syllable.

(97) /fihm-na/ > ['fihimna] 'our understanding'
/ʔakl-ha/ > ['ʔakilha] 'her food'

(from Brame 1974)

Prominence in forms such as these is opaque, i.e. it does not agree with surface syllabification. In DT, such a situation can be explained by making reference to an intermediate level of representation, at which epenthesis has not yet applied. This can be achieved simply by ordering stress assignment before epenthesis, so that the epenthetic vowel is not present at the moment in which prominence is assigned, and cannot therefore be stressed (Kenstowicz 1981, Brame 1974). This analysis would also account for forms such as Iraqi Arabic *kitabit* [ki'tabit] '(I) wrote', in which the epenthetic vowel seems to be invisible to the stress assignment algorithm (see 1.4). If such a vowel were taken into account, the second syllable would be light, and stress would fall on the first syllable, yielding the unattested *['kitabit] (Broselow 1992:27).

In OT, which involves parallel evaluation of constraints rather than serial application of rules or processes, no intermediate levels are postulated between input and output. Moreover, since it is assumed that input strings are metrically unparsed, stress assignment cannot be due to input-output correspondence between stressed segments. Alderete (1995), however, puts forward a head dependence constraint that establishes a correspondence between stressed output vowels and input vowels at large. Such a constraint is called HEADDEP-IO, and states that 'every vowel in the output's prosodic head has a correspondent in the input'. Since epenthetic vowels lack a correspondent in the input, they cannot belong to a prosodic head, i.e. they cannot be stressed. In these cases, HEADDEP-IO dominates the well-formedness constraint WSP (Weight-to-Stress-Principle), requiring that heavy syllables must be stressed. Making use of this ranking and of the OO-identity constraint HEADMAX-BA (see ch. 4), Kager (1999:287-92) accounts for the evaluation of a form like /fihm-na/ > ['fihimna] 'our understanding' as in the following tableau.

(98)

Input: /fihm-na/ Base: ['fi.him]	*COMPLEX	HEAD MAX-BA	NO-i]$_\sigma$	HEAD DEP-IO	MAX-BA	WSP
a. [fi.'him.na]			*	*!		
b. ☞ ['fi.him.na]			*			*
c. ['fih.mi.na]			*		*!	
d. ['fhim.na]		*!			*	
e. ['fihm.na]	*!				*	

(from Kager 1999:292)

In tableau (98), while the candidate which is faithful to the input (i.e. candidate e.) is eliminated by *COMPLEX because it contains a coda cluster, HEADDEP-IO rules out candidate a., in which stress is assigned to the epenthetic vowel. HEADMAX-BA, requiring vowels which are stressed in the base to have correspondents in the affixed form, is violated by candidate d. This is because the first [i], whose correspondent in the base is stressed, is deleted. Finally, in order for the candidate with the wrong insertion site not to be selected, MAX-BA is made to outrank WSP. As was explained above, the same result could be achieved by means of the constraint ALIGN-L(σ, PW).[41]

In a dialect like AA, however, epenthetic vowels seem to be systematically stressed when the stress assignment algorithm requires it (see (99)).

(99)	/ʔul-t-l-ha/	>	[ʔul'tilha] 'I said to her'
	/katab-t-l-na/	>	[katab'tilna] 'you (m.sg.) wrote to us'
	/katab-t-l-kum/	>	[katab'tilkum] '(I) wrote to you (pl.)'
	/ʔul-t-l-u/	>	[ʔul'tillu] 'I said to him'
	/ʒa:b-l-ha/	>	[ʒa:'bilha] 'he brought to her'
	/fihm-na/	>	[fi'himna] 'our understanding'
	/akl-ha/	>	[ʔa'kilha] 'her food'

In a derivational framework, in this dialect epenthesis must be ordered before stress assignment; in an OT perspective, WSP must outrank HEADDEP-IO.

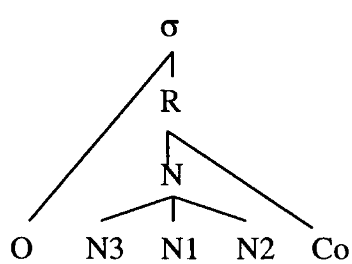
[41] In Kiparsky's (2000) LPM-OT analysis (see 4.3), the reason why epenthesis is invisible to stress assignment in dialects like Palestinian Arabic is that, while VI is a post-lexical process, word-prominence is assigned in the lexicon. Since the input for post-lexical evaluation is the output of lexical evaluation, IO-faithfulness to stress, when ranked high, preserves the location of lexical prominence. It is a more sophisticated way of saying, with Brame (1974), that epenthesis is ordered after stress assignment.

CHAPTER 6: CONCLUSION

6.1 Vocoids and their distribution

This thesis has shown that vocoids can be construed as a well-defined class of sounds not only from a phonetic, but also from a phonological point of view. Such a class comprises vowels and glides, or semi-vowels, that is all those segments which are produced with open approximation and may form syllabic nuclei in all natural languages. They share a high degree of openness, intensity and periodicity, i.e., in phonological terms, of sonority. Although all vocoids are sonorant sounds, the feature that characterizes them exclusively as a category is [-consonantal]. Therefore, vocoids may be construed as a class on segmental grounds: a semi-vowel exhibits exactly the same feature matrix as the homorganic vowel, while the difference between these two subclasses is given merely by their syllabic distribution. Since syllabification is not generally present at the level of deep structure, there is no distinction between glides and vowels underlyingly. When syllabic structure is assigned, certain vocoids may surface either as glides (if they are syllabified in a marginal position such as onset or coda) or as vowels (if they are syllabified in the syllable peak). In reality, up to five positions can be distinguished within the syllable, all of which may be occupied by vocoids. Apart from the traditional onset (O), peak (N1) and coda (Co), a post-peak nuclear node (N2) and a pre-peak nuclear node (N3) have been identified (see diagram in (1)). The former is the typical position that a post-vocalic glide, or 'off-glide', occupies (Co-glides are less common cross-linguistically), as in English *howl* [haṷl]. An N3-vocoid appears instead in certain rising diphthongs, such as Italian 'dittonghi mobili' (e.g. *uomo* [wɔmo] 'man', *piede* [pjɛde] 'foot').

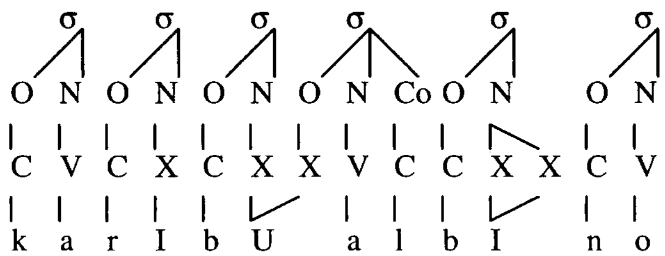
(1)



Probably the strongest argument for circumscribing a class of vocoids including both vowels and glides is that the same segment (generally a high vocoid) may surface as a glide or as a vowel according to the syllabic position to which it is assigned. For instance, the high vocoid /I/ in a stem like Arabic /ɣa:lI-/ 'expensive' surfaces as a vowel when word-final (cf. [ɣa:**li**] 'expensive (m.sg.)'), as a glide when followed by another vocoid (cf. [ɣa:**lj**a] 'expensive (f.sg.)').

Another case in point is so-called Glide Insertion (see ch. 5), by which a vowel lengthens to fill in an adjacent empty onset. In the Italian phrase *caribú albino* [ka.ri.bu.wal.bi:.no] 'albino caribou', the features of the final vocoid of *caribú* spread into the empty onset of the following syllable [_al]$_\sigma$, as illustrated in (2).

(2)



In these instances, the same vocoid occupies two (or three) positions at the same time: one (or two, if the vocoid is already long) nuclear position and one onset position.

Some vocoids, however, never alternate in their syllabicity, even when the phonological environment would seem to require it. The simplest case is that of highly sonorous vocoids, which are generally barred from marginal positions in the syllable. In both Italian and Arabic, for instance, only high vocoids may alternate in syllabicity, while mid and low vocoids always surface as vowels. The final vocoid in an Italian word like *carro* 'chariot', not being high, never yields a glide when followed by another vocoid, as in *carro armato* 'tank': it is either deleted or it preserves its nuclear position, forming a hiatus with the following vowel. Even more interestingly, some high vocoids are systematically realized as a vowel or as a glide, regardless of the environment (see ch. 2). In Temiar, for instance, certain vocoids are always non-syllabic, even when they are surrounded by consonants, as in /kwkɔ:w/ > [kewkɔ:w] (not *[kukɔ:w]) 'call (active, continuative)'. In Ait Seghrouchen Berber, on the other hand, there are vocoids that always surface as vowels, requiring j-insertion (see the minimal pair *tessu(j)ax* 'she made us a bed' ~ *tesswax* 'she made us drink'). A similar case is that of Italian words like *fantasia* [fan.ta.'zi.a] 'imagination', in which the high vocoid is realized as a vowel even if it is adjacent to another vocoid. Since, in this language, syllabification must precede stress assignment, it cannot be the presence of prominence on the target vocoid that blocks Glide Formation.

Two separate approaches to deal with non-alternating glides and non-alternating vowels were proposed (and rejected) in chapter 2. Non-alternating glides could be seen as consonantal segments, as opposed to alternating, [-consonantal] vocoids. Non-alternating vowels could instead be attached to syllable heads underlyingly. Among other drawbacks, however, neither of these approaches can account for both subclasses of segments. A unified account is instead made possible by postulating a CVX-tier, i.e. a skeleton on which timing slots are of three types: V-slots, which may only attach to nuclear positions, C-slots, which only attach to marginal subsyllabic nodes and X-slots, projected by alternating vocoids. In this perspective, the above-mentioned Temiar glides will be linked to C-slots, while non-alternating vowels in Berber or Italian will be linked to V-slots.

6.2 Prosodic structure

6.2.1 Syllabification

As is clear from the above-mentioned considerations, the distribution of vocoids is strictly connected with syllable structure. This is not surprising, since, in many languages, only vocoids may occupy syllable nuclei. Hence, if no vocoid is present, no syllable can be formed. Phenomena like Vowel Deletion (VD) and Glide Formation (GF), by eliminating potentially syllabic elements or making them non-syllabic, change the prosodic make-up of words. A process like Vowel Insertion (VI), on the other hand, builds new syllabic structure.

In a more sophisticated framework, the phenomena that have been analyzed in this thesis can almost always be seen as a direct result of syllabification (or resyllabification). A templatic approach postulating directional parsing of phonological material has proved to be particularly useful in this respect, in that many processes affecting vocoids can be made to fall out naturally from the matching of syllabic templates with the segmental string. That is, as syllabification proceeds, vocoids automatically assume different shapes according to the syllabic position they are assigned. The syllabification of a high vocoid as a glide or as a vowel thus follows directly from its location with respect to other segments: if it is adjacent to a V-slot element, it is syllabified in a marginal position, while if it is surrounded by consonantal sounds it goes to occupy a syllable nucleus. A vocoid is instead deleted when it is skipped by the syllabification mechanism, due to a mismatch between segmental and syllabic structure. VI applies when a nucleus-less syllable is formed to incorporate consonants that would otherwise be left unparsed.

A directional approach, however, cannot always portray phonological operations in a straightforward fashion. Generally, other conditions independent of directional syllabification must be stipulated in order to obtain the correct surface forms. Moreover, the mechanism often has to look ahead to see if its output will be syllabifiable. For instance, in a left-to-right language with obligatory onsets, in order to syllabify an alternating vocoid correctly the mechanism must know if the next segment is a vowel or a consonant: if it is a vowel, the vocoid will be syllabified with it as an onset, if it is a consonant, the vocoid will occupy a nucleus (or coda) position.

A more fundamental fault with this framework is that opposite directions of syllabification are sometimes required to account for different phenomena within the same linguistic variety. For example, Arabic dialects should all have left-to-right syllabification in order to favour coda formation over onset maximization. However, epenthesis and syncope in some of these varieties pattern in a way that can only be explained by postulating right-to-left syllabification. Moreover, divergence between syncope and epenthesis outputs within the same dialect can be found. In AA, for instance, right-to-left syllabification accounts for forms such as *Hajarikbiir* '(a) big stone', but obtains the wrong output for forms like *bjuktubu* '(they) write', which can instead be derived by left-to-right parsing (see 5.3).

In addition, in order to obtain the correct result for syncope in some Arabic dialects, all potentially syncopated vowels must be absent from the UR. This assumption has two main drawbacks, one of which is that, if all vowels alternating with zero are epenthetic, differences in the quality of such vowels cannot be accounted for. The other drawback is that potentially syncopated vowels may behave differently from epenthetic vowels, for example with respect to stress assignment.

6.2.2 Foot structure

Foot structure may also affect the distribution of vocoids. Foot (and word) minimality often triggers quantitative alternations in these segments. For instance, vowels may be lengthened in order to avoid subminimal monomoraic feet. For analogous reasons, in a weight-insensitive language, a vowel may be inserted to make a foot disyllabic. The insertion of a vowel may actually be preferred over the deletion of a consonant in order to avoid subminimal feet. In Tuscan Italian, for example, the non-syllabifiable final consonant of a word like /tram/ 'tram' triggers epenthesis, as in [tram.m**e**], rather than consonant deletion, as in the subminimal *[tra] (see 5.2). VD may also be blocked by foot minimality. In an Italian sentence such as *pianterá abeti* '(s/he) will plant fir trees', the two clashing low vocoids contract into a long vowel ([(pjan.te.)(r**a:**.)(be:.ti)]) rather than degeminating (*[(pjan.te.)(r**a**.)(be:.ti)]), because this latter process would yield a monomoraic foot *$(\text{ra})_F$.

On the other hand, since binary feet are generally preferred over ternary feet, a vowel is often deleted to eliminate one of two consecutive unstressed syllables, as in English *memory* ['mɛm.ɹi]. Also, because of the tendency to maintain the temporal distance between two stresses constant (foot isochrony), vowels in monosyllabic feet tend to be longer than vowels in polysyllabic feet. A consequence of foot isochrony is, for example, Trisyllabic Shortening, by which the stressed vowel in trisyllabic feet is shorter than that in monosyllabic and disyllabic ones (cf. English *profane* [prɒ'fe̯ɪn] ~ *profanity* [prɒ'f**æ**nɪti]).

The consideration that the optimal iamb tends to be uneven (i.e. LH), while the optimal trochee tends to be even (**LL** or **H**) may also affect the distribution of vocoids. For instance, the vowel of the second syllable of an iambic foot may be lengthened in order to obtain an uneven iamb. On the other hand, a vowel may be shortened in order to yield an even trochee (see English *cone* [kə̯ʊn] ~ *conic* [kɒ.nik]).

Foot well-formedness constraints may trigger truncation, too. In Italian, for instance, this process applies to obtain an even trochee (**H**) out of an uneven one (**HL**), as in *mare* [ma:re] > *mar* [mar] 'sea' (see 4.2).

6.2.3 Stress assignment

Since the syllable nucleus is the stress-bearing unit, and vocoids are typically nuclear elements, stress is often relevant to phenomena involving this class of sounds.

Stressed vowels are generally more resistant to reduction than unstressed ones. In many languages, for instance, vowels bearing prominence do not alternate with glides. GF may also be blocked by stress on the second clashing vocoid, as in Italian *ti ámo* [ti'(j)a:mo] '(I) love you'. In addition, stressed vocoids are more resistant to shortening, and are rarely deleted. In AA, for example, only unstressed vowels undergo syncope, as in *fihímti* ['fhimti] '(you f.sg.) understood'; when the same vowel is in a position to be assigned prominence, it is preserved (cf. *fíhim* ['fihim] '(he) understood').

On the other hand, stressed syllables - and thus vocoids - are often lengthened. In quantity-sensitive languages, there is a close relationship between syllable weight and stress assignment. Since a long vowel makes a syllable heavy, vowel length is also related to prominence (see ch. 3). Apart from the fact that stressed vowels are always phonetically longer than unstressed ones, in some languages stress triggers vowel lengthening. In Italian, for instance, stressed syllables must be bimoraic; therefore, vowels in open stressed syllables are lengthened (e.g. *fáro* ['fa:.ro] 'lighthouse'). In other languages, in which vowel length is underlying, long vowels attract stress. This is the case in many Arabic dialects, in which the rightmost heavy syllable is assigned prominence (e.g. *majnúun* 'crazy', *tháani* 'second').

6.3 Vocoids in Italian and Arabic

This thesis makes special reference to two languages, Standard Italian (SI) and Ammani Arabic (AA). In their treatment of vocoids, some general trends can be detected, which may be attributed to universal principles or constraints. In both languages, vocoids alternate in syllabicity, and syllabification is the main trigger for processes such as deletions, insertions and length alternations. However, divergences in the morphological role of vocoids and in prosodic structure lead to differences in the implementation of such processes. For instance, while both languages exhibit glides and vowels, only Arabic has coda vocoids.

Both in Arabic and in Italian, high vowels alternate with on- and off-glides, although GF is more common in the latter language, while in the former it is often replaced by ʔ-insertion (see ch. 2). In Arabic, which has a non-concatenative morphology, alternations take place on both root and affix tiers. This is because such tiers are morphological, rather than phonological, entities.

In Italian, however, some high vocoids must be underlyingly pre-linked to a syllable head, since they surface in a peak position even when the syllabification algorithm would predict otherwise (e.g. *Maria* [ma.ri:.a], and not *[ma:.rja]). In Arabic, on the other hand, the distribution of glides and vowels seems to be entirely predictable according to morphological and prosodic information. Nevertheless, the fact that level 1 affixes are entirely made up of vocoids and roots only exhibit segments that can occupy marginal positions is suspicious. Further study may demonstrate that root and affixal vocoids have to be distinguished phonologically.

As far as vowel quantity is concerned, Arabic has phonological length distinctions, while in Italian vowel-length is phonetic (see ch. 3). Syllables are maximally and optimally bimoraic in both languages, but SI seems to implement this principle more strictly than AA. In Italian trimoraic syllables are completely banned, so that vowels in closed syllables are always short. AA, on the other hand, exceptionally allows even word-internal CVVC syllables.[1] Moreover, in SI stressed syllables are not only maximally, but also minimally bimoraic, so that a vowel in an open stressed syllable is always long. In Arabic, on the contrary, monomoraic syllables may bear prominence (e.g. *balad* ['ba.lad] 'village; country'). Although in both languages heavy syllables attract stress, the relation between syllable weight and stress assignment differs in some respects. While in Italian word prominence is distinctive, and often makes syllables heavy by lengthening their vowels, in AA stress assignment is entirely determined by syllable weight - and thus, to a certain extent, by vowel length.

In SI, VD is much less common than in the Arabic dialects, and is generally limited to the elision of certain pre-vocalic non-high vocoids. Syncope only applies in fast and careless speech. In Arabic, on the other hand, vowel syncope is the most common deletion phenomenon (see ch. 4).

While VI is very common in the Arabic dialects, it is virtually absent from SI, and can only be found in non-standard varieties of the Romance language (see ch. 5). Both in non-standard Italian and in Arabic, epenthesis applies to rescue non-syllabifiable consonants from erasure. However, the two languages exhibit different syllable structure constraints: while Italian maximizes onsets but has a modified Coda Filter by which only certain consonants are allowed in syllable codas, Arabic tends to allow

[1] However, as was explained in chapter 3, these may be construed as bimoraic by means of Broselow's (1992) 'Adjunction to Mora'.

only simple onsets and codas, but has no Coda Filter. Moreover, onsets are obligatory in the latter language, but not in Italian. These differences have a bearing on the implementation of epenthesis. For instance, a vowel may be inserted after a single coda consonant in Italian, but not in Arabic. This latter language, on the other hand, may break up onset clusters which are instead allowed in Italian.

It should also be pointed out that Consonant Deletion is a more common alternative to VI in Italian than it is in Arabic. In this language, the tendency to preserve consonants may be related to the fact that these segments often make up the root melody: their deletion would have more serious semantic implications than that of affixal vocoids, which are in fact syncopated (and inserted) rather freely. The fact that VD and VI processes apply in contrasting environments in different Arabic dialects also seems to be linked to the lower significance of affixal segments.

More generally, vowels may have a different status in Italian and Arabic not only from a morphological, but also from a phonological point of view. In the Romance language, the distribution of vocoids is generally more rigid and less affected by processes such as deletion or insertion. Due to their lower morphological significance, coupled with the fact that short vowels are not represented in the spelling, Arabic vocoids follow more variable patterns. The fact that SI has more vowel phonemes and AA more allophones is perhaps related to these considerations.

6.4 Optimality Theory and Derivational Theory

In this thesis, the analysis is generally couched in a theory that comprises autosegmental representations, prosodic structure, and a multi-stratal rule-based derivational framework. This type of approach, which I have called Derivational Theory (DT), has been contrasted with a constraint-based, output driven approach which does not allow intermediate levels of representation between input and output forms, Optimality Theory (OT). Most of the phenomena that were analyzed in this thesis can be accounted for in a rather satisfactory way in both frameworks, although specific instances may be characterized more economically and explanatorily within one or the other approach.

The main advantage of OT is that it renders prosodic processes such as glide ~ vowel alternation, length alternations, deletion and epenthesis transparent, in that it focusses on their output targets rather than on the change from input to output. Such processes are generally aimed at achieving the well-formedness of prosodic constituents such as the syllable or the foot. In traditional DT, well-formedness is realized first in metrical parsing, then in the above-mentioned quantitative operations, without making the connection explicit. Kager (1999:186) points out that such an analysis renders processes like syncope and epenthesis opaque, in that it has to postulate an intermediate level at which segments are exhaustively parsed into syllables and syllables into feet, thus obscuring the conditioning environment. In other words, it misses the generalization that these phenomena conspire toward the implementation of syllable and foot well-formedness. Moreover, DT often has to resort to prosodic structure constraints anyway, such as the Bimoraicity Constraint, the Onset Principle, or the Coda Filter. For instance, a process like GF will be formulated as a rule delinking a high vocoid from a nuclear position and resyllabifying it into the following onset. In order to motivate such an operation, an Onset Principle – absolute in cases like Arabic, preferential in cases like Italian – may be postulated.

In OT, on the other hand, the distinction between initial metrical parsing and later quantitative processes disappears, since well-formedness is evaluated at the output level (Kager 1999:175). A phenomenon like GF can thus be accounted for by means of the interaction of markedness constraints such as NoHiatus and *M/V (see ch. 2).

More innovative derivational frameworks, such as the directional approach and other prosodic models, make most processes of this kind fall out naturally from syllabification principles, without any need for creating or deleting metrical structure. However, even in these approaches certain prohibitions or preferences must be stipulated independently of the parsing mechanism. In SI, for instance, sonority violating domain-initial consonants are systematically preserved, and later incorporated into the following syllable, rather than undergoing deletion (see 5.2). There is no way in which this preference can be expressed as an integral part of syllabification: it must be stipulated on the side. In OT, on the contrary, prohibitions and preferences are stated straightforwardly by means of constraints which are the basis of the whole theory itself. The fact that such constraints are universal but violable captures the generalization that principles such as Bimoraicity, Sonority Sequencing or the Onset Principle are present in all languages, but are implemented with different degrees of strength.

On the whole, therefore, OT seems to offer valuable insights into the complexities of phonological phenomena. However, this approach also has several drawbacks with respect to DT frameworks. Two main problems, in particular, have not as yet found satisfactory solutions within the theory.

The first problem is to do with the complete denial of intermediate levels of representation between input and output. Certain phenomena, in fact, can only be accounted for by introducing a multi-stratal model. Probably the clearest, and the most studied, of such phenomena is opacity, by which output alternations are based on generalizations which are not apparent on the surface. A typical example of an opaque process is Compensatory Lengthening, i.e. the lengthening of a vowel as a consequence of the deletion of the following consonant (see ch. 3). Since the deletion of the consonant constitutes the triggering environment for the lengthening of the vowel, the two processes cannot apply in parallel. Consequently, at least one intermediate stage in the derivation must be postulated, at which deletion has taken place but lengthening has not.

In order to account for opaque phenomena without renouncing OT's parallelism, various models have been proposed, such as Sympathy or input-output markedness. However, the most satisfactory solution seems to be the actual introduction of intermediate levels of representation into the theory, as in Kiparsky's (2000) LPM-OT. As was shown in 4.2, this model is also useful in cases in which the preference for different repair strategies within and outside a certain domain cannot be accounted for on the basis of alignment or anchoring constraints. Although Kiparsky's (2000) proposal may seem to somewhat weaken the theory, a constraint-based derivational framework is not necessarily a contradiction, and may offer significant insights into phonological phenomena.

The second major problem with a constraint-based approach is the lack of definition of the lexical input. This aspect has been neglected by OT phonologists on the basis of the assumption that no restrictions hold at the underlying level (Richness of the Base): whatever the input, output constraints will always yield the correct surface forms. This indeterminacy, however, is in clear contrast with the requirements of IO-faithfulness constraints such as MAX-IO or DEP-IO, which demand correspondence between an output form and a well-defined input. Some correspondence constraints, such as MAX-μ-IO, even seem to assume that a certain amount of prosodic structure is already assigned in the input. Moreover, lexically marked exceptions, such as non-alternating vocoids, are problematic for an approach assuming Richness of the Base.

Further study may reveal that, by introducing a sub-theory of inputs (perhaps comprising a certain amount of underspecification), one may achieve a better characterization of linguistic phenomena.

More specific problems concern the formulation of individual OT constraints. While many constraints, and especially those concerning prosodic markedness, translate universal tendencies in a perfectly transparent fashion, others are ad hoc and lack phonetic grounding (see for example a featural markedness constraints such as *[-back]). Besides, the implementation of certain restrictions is rather convoluted, as in the case of gradient alignment constraints. Finally, constraints which have been postulated to account for specific phenomena may put forward opposite requirements to other constraints, again raising the question of phonetic grounding. Examples of such pairs of constraints, which clearly reflect parametrical choices, are *[-high] / *[+high], RHTYPE=I / RHTYPE=T, ALIGN-L(σ, PW) / ALIGN-R(σ, PW). In these cases, since the action of the higher-ranked constraint completely obscures that of the lower-ranked constraint, some phonologists seem to assume that only one of them is present in a given language. This assumption, however, undermines the concept of universality which is at the basis of OT.

6.5 Suggestions for further research

While this thesis has illustrated some of the trends that characterize the class of vocoids, it only makes reference to two linguistic varieties and to a handful of phonological processes. Further study of prosodic alternations involving vocoids in other languages is therefore required in order to detect more general patterns and to confirm or refute the conclusions I have reached.

Moreover, since the two theoretical approaches adopted in this thesis each have advantages and disadvantages, they should perhaps be integrated into a third framework. Further research may reveal that a constraint-based derivational theory establishes a more complete framework within which phonological phenomena can be analyzed.

The status of Arabic affixal vocoids should also be further investigated.

REFERENCES

Abdo, D.A. (1969) 'Stress and Arabic Phonology'. Ph.D. dissertation, University of Illinois, Urbana.

Abercrombie, D. (1967) *Elements of General Phonetics*. Chicago: Aldine.

Abu-Mansour, M.H. (1991) Epenthesis in Makkan Arabic: unsyllabified consonants vs. degenerate syllables. In B. Comrie and M. Eid (eds.) *Perspectives on Arabic Linguistics III.* Amsterdam/Philadelphia: John Benjamins, pp. 137-154.

Abu-Mansour, M.H. (1992) Closed syllable shortening and morphological levels. In E. Broselow, M. Eid and J. Mc.Carthy (eds.) *Perspectives on Arabic Linguistics IV.* Amsterdam/Philadelphia: John Benjamins, pp. 47-75.

Abu-Mansour, M.H. (1995) Optimality and conspiracy in the syllable structure of Arabic. *University of Massachusetts Occasional Papers 18*, pp. 1-20.

Abu-Salim, I. (1982) 'A Reanalysis of Some Aspects of Arabic Phonology: a metrical approach'. Ph.D. dissertation, University of Illinois, Urbana. (in Younes 1995)

Abumdas, A.H.A. (1985) 'Libyan Arabic Phonology'. Ph.D. dissertation, University of Michigan. Distributed by University Microfilms International, Ann Arbor, Michigan, U.S.A.

Al-Ageli, H. (1995) 'Syllabic and Metrical Structure in Tripolitanian Arabic'. Ph.D. thesis, University of Essex, Colchester.

Al-Ani, S.H. (1970) *Arabic Phonology: an Acoustical and Physiological Investigation.* The Hague/ Paris: Mouton.

Alderete, J. (1995) Faithfulness to prosodic heads. Ms., University of Massachusetts, Amherst.

Andersen, H. (1972) Diphthongisation. *Language* 48/1, pp. 11-47.

Anderson, S.R. (1984) A metrical interpretation of some traditional claims about quantity and stress. In M. Aronoff and R.T. Oehrle (eds.) *Language Sound Structure: Studies in Phonology Presented to Morris Halle by his Teacher and Students.* Cambridge, MA: MIT Press, pp. 83-106.

Angoujard, J.P. (1990) *Metrical Structure of Arabic.* Dordrecht: Foris Publications.

Archangeli, D. (1984) 'Underspecification in Yawelmani Phonology and Morphology'. Ph.D. dissertation, M.I.T., Cambridge, MA.

Bauer, L. (1984) Linking /r/ in RP: some facts. *Journal of the International Phonetic Association* 14/2, pp. 74-79.

Beckman, M.E. and I. Edwards (1990) 'Lengthenings and shortenings and the nature of prosodic constituency'. In J. Kingston and M.E. Beckman (eds.) *Papers in Laboratory Phonology I,* pp. 152-178.

Benhallam, A. (1980) 'Syllable Structure and Rule Types in Arabic'. PhD thesis, University of Florida.

Berendsen, E. and O.E. den Os (1987) Glide insertion: domains, speech rate and phonetic prominence. In F.Beukema and P.Coopmans (eds.) *Linguistics in the Netherlands 1987.* AVT publications 3, Dordrecht: Foris, pp. 13-20.

Booij, G. (1988) Review of Nespor and Vogel (1986). *Journal of Linguistics 24*, pp. 515-25.

Booij, G. and R. Lieber (1993) On the simultaneity of morphological and prosodic structure. In S. Hargus and E. Kaisse (eds.) *Studies in Lexical Phonology.* San Diego: Academic Press, pp. 23-44.

Bosisio, N. (1998) 'Glide Formation in Italian and Glide Insertion in English: a Prosodic Phonology Perspective'. MA dissertation, University of Durham, UK.

Brame, M.K. (1970) 'Arabic Phonology: Implications for Phonological Theory and Historical Semitic'. Ph.D. dissertation, MIT, Cambridge, Massachusetts.

Brame, M.K. (1972) The segmental cycle. In M.K. Brame (ed.) *Contributions to Generative Phonology.* Austin/ London:Uniersity of Texas Press, pp. 62-72.

Brame, M.K. (1974) The cycle in phonology: stress in Palestinian, Maltese and Spanish. *Linguistic Inquiry 5*, pp. 39-60.

Broselow, E. (1979) Cairene Arabic syllable structure. *Linguistic Analysis 5*, pp. 345-382.

Broselow, E. (1992) Parametric variation in Arabic dialect phonology. In E. Broselow, M. Eid and J. Mc.Carthy (eds.) *Perspectives on Arabic Linguistics IV.* Amsterdam/Philadelphia: John Benjamins, pp. 7-45.

Broselow, E., M.Huffman, S.Chen and R.Hsieh (1995) The timing structure of CVVC syllables. In M.Eid (ed.) *Perspectives on Arabic Linguistics VII.* Amsterdam/Philadelphia: John Benjamins, pp. 119-138.

Broselow, E., S.Chen and M.Huffman (1997) Syllable weight: convergence of phonology and phonetics. *Phonology* 14/1, pp. 47-82.

Brown, A. (1988) Linking, intrusive and rhotic /r/ in pronunciation models. *Journal of the International Phonetic Association* 18/2, pp.144-151.

Bullock, B.E. (1991) 'The Mora and the Syllable as Prosodic Licensers in the Lexicon'. Ph.D. dissertation, University of Delaware.

Bullock, B.E. (1998) The myth of equivalence: where two lights do not make a long. In J. Lema and E. Treviño (eds.) *Theoretical Analyses on Romance Languages.* Amsterdam/ Philadelphia: John Benjamins.

Burzio, L. (1987) Prosodic reduction. *Linguistic Symposium on Romance Languages XVII,* Ohio State University.

Burzio, L. (1994) *Principles of English Stress.* Cambridge: CUP.

Cantineau, J. (1960) *Etudes de Linguistique Arabe.* Paris: Librairie C. Klincksieck.

Carr, P. (1993) *Phonology.* Basingstoke/ London: Macmillan.

Catford, J.C. (1977) Mountain of tongues: the languages of the Caucasus. *Annual Review of Anthropology 6,* 283-314. (in Ladefoged and Maddieson 1996)

Chierchia, G. (1986) Length, syllabification and the phonological cycle in Italian. *Journal of Italian Linguistics* 8/1, pp. 5-34.

Chitoran, I. (1997) 'The Phonology and Morphology of Romanian Glides and Diphthongs: a Constraint-Based Approach'. PhD dissertation, Cornell University, in *Dissertation Abstracts International* 58/5, p. 1680.

Chomsky, N. (1957) *Syntactic Structures.* The Hague: Mouton.

Chomsky, N. (1965) *Aspects of the Theory of Syntax.* Cambridge, MA: MIT press.

Chomsky, N. (1981) *Lectures on Government and Binding.* Dordrecht: Foris.

Chomsky, N. and M. Halle (1968) *The Sound Pattern of English (SPE).* New York: Harper and Row.

Clements, G.N. and S.J. Keyser (1983) *CV Phonology: a Generative Theory of the Syllable.* Cambridge, Mass.: MIT press.

De Chene, B.E. (1985) *The Historical Phonology of Vowel Length.* New York/ London: Garland Publishing.

Dell, F. and M. Elmedlaoui (1985) Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics 7,* pp. 105-30.

Derwing, B.L., D.B.Parkinson and R.A.Beinert (1995) Experimental investigations of Arabic syllable structure. In M.Eid (ed.) *Perspectives on Arabic Linguistics VII.* Amsterdam/Philadelphia: John Benjamins, pp. 107-118.

Dickins, J. (1996) 'Some speculations on the vowel system of Classical Arabic'. *CMEIS Occasional Paper nr 53,* University of Durham, pp. 6-19.

Di Luzio, A. (1988) On some segmental phonological processes and their suprasegmental conditioning in Standard Italian and in an Abruzzian dialect. In P.M.Bertinetto and M. Loporcaro (eds.) *Certamen Phonologicum: Papers from the 1987 Cortona Phonology Meeting.* Torino: Rosenberg and Sellier, pp. 299-335.

Dixon, R.M.W. (1977a) Some phonological rules of Yidin^y. *Linguistic Inquiry 8,* pp. 1-34. (in Nespor and Vogel 1986)

Dixon, R.M.W. (1977b) *A Grammar of Yidin^y*. Cambridge: CUP. (in Nespor and Vogel 1986)

Donegan, P.J. (1985) *On the Natural Phonology of Vowels.* New York and London:Garland Publishing.

Durand, J. (1987) On the phonological status of glides: the evidence from Malay. In J. Anderson and J. Durand (eds.) *Explorations in Dependency Phonology.* Dordrecht: Foris, pp. 79-107.

Fudge, E. (1969) Syllables. *Journal of Linguistics 5,* pp. 253-87.

Fudge, E. (1999) Words and feet. *Journal of Linguistics* 35/2, pp. 273-296.

Gieser, C. (1970) The morphophonemic system of Guininaang (Kalinga). *Philippine Journal of Linguistics 1,* pp. 52-68. (in Kenstowicz and Kisseberth 1979)

Gimson, A.C. (1989) *An Introduction to the Pronunciation of English.* London: Arnold.

Goldsmith, J. (1976) 'Autosegmental Phonology'. Ph.D. thesis, MIT, Cambridge, Mass. Distributed by Indiana University Linguistic Club. Published in 1979 by Garland Press, New York.

Grammont, M. (1933) *Traité de Phonétique.* Paris: Delagrave.

Guerrsel, M. (1986) Glides in Berber and syllabicity. *Linguistic Inquiry 17,* pp. 1-12.

Halle, M. (1992) Phonological features. In W. Bright (ed.) *International encyclopedia of linguistics,* vol. 3. Oxford: OUP, pp. 207-12.

Halle, M. and W.J. Idsardi (1995) General properties of stress and metrical structure. In J.A. Goldsmith (ed.) *The Handbook of Phonological Theory.* Oxford: Blackwell, pp. 403-43.

Halle, M. and J.R. Vergnaud (1980) Three-dimensional phonology. *Journal of Linguistic Research 1*, pp. 83-105.

Halle, M. and J.R. Vergnaud (1987) *An Essay on Stress.* Cambridge, Mass.: MIT Press.

Hammond, M. (1997) Optimality theory and prosody. Ms., University of Arizona.

Hannahs, S.J. (1995) The phonological word in French. *Linguistics* 33, pp. 1125-1144.

Hannahs, S.J. (1998a) Unexceptional exceptions and French glides. *LSRL* 28, Penn State University, May 1998. Ms., University of Durham.

Hannahs, S.J. (1998b) *Characterising French Glides.* Ms., University of Durham.

Haraguchi, S. (1984) Some tonal and segmental effects of vowel height in Japanese. In M. Aronoff and R.T. Oehrle (eds.) *Language Sound Structure: Studies in Phonology Presented to Morris Halle by his Teacher and Students.* Cambridge, MA: MIT Press, pp. 145-156.

Harms, R. (1968) *Introduction to Phonological Theory.* Englewood Cliffs, N.J.: Prentice-Hall. (in Kenstowicz 1994)

Harrell, R.S. (1957) *The Phonology of Colloquial Egyptian Arabic.* New York: American Council of Learned Societies.

Harris, J. (1983) *Syllable Structure and Stress in Spanish: a non-linear analysis.* Cambridge, Mass.: MIT Press.

Harris, J. (1994) *English Sound Structure.* Oxford: Blackwell.

Harris, J. and E.M. Kaisse (1999) Palatal vowels, glides and obstruents in Argentinian Spanish. *Phonology 16/2,* pp. 117-90.

Hayes, B. (1980) 'A Metrical Theory of Stress Rules'. Ph.D. dissertation, MIT, Cambridge, Massachusetts.

Hayes, B. (1989) Compensatory lengthening in moraic phonology. *Linguistic Inquiry 20*, pp. 253-306.

Hayes, B. (1990) Precompiled phrasal phonology. In S. Inkelas and D. Zec (eds.) *The Phonology-Syntax Connection.* Chicago: Chicago University Press, pp. 85-108.

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

Hieke, A.E. (1984) Linking as a marker of fluent speech. *Language and Speech* 27/4, pp. 343-354.

Holes, C. (1995) *Modern Arabic: Structures, Functions and Varieties.* London: Longman.

Hyman, L. (1985) *A Theory of Phonological Weight.* Dordrecht: Foris.

Idsardi, W. (1992) 'The Computation of Prosody'. Ph.D. dissertation, MIT, Cambridge, Massachusetts

Inkelas, S. (1989) 'Prosodic Constituency in the Lexicon'. Ph.D. dissertation, Stanford University.

Inkelas, S. (1995) The consequences of optimization for underspecification. In J. Beckman (ed.) *Proceedings of the North East Linguistic Society 25*, pp. 287-302.

Itô, J. (1986) 'Syllable Theory in Prosodic Phonology'. Ph.D. dissertation, University of Massachusetts, Amherst. Published by Garland Press, New York, 1988.

Itô, J. (1989) A prosodic theory of epenthesis. *Natural Language and Linguistic Theory* 7, pp. 217-259.

Itô, J. and A. Mester (1992) Weak layering and word binarity. Ms., University of California, Santa Cruz.

Itô, J., A. Mester and J. Padgett (1995) Licensing and underspecification in Optimality Theory. *Linguistic Inquiry 26*, 571-614.

Jakobson, R. (1928) Proposition au Premier Congrès International de Linguistes. Quelles sont les méthodes les mieux appropriées à un exposé complet et pratique de la phonologie d'une langue quelquonque? In R. Jakobson (1962) *Selected Writings,* vol. 1. The Hague: Mouton.

Jakobson R., G. Fant and M. Halle (1952) *Preliminaries to Speech Analysis.* Cambridge, Mass.: MIT press.

James, A.R. (1988) Review article on M.Nespor and I.Vogel (1986). *Prosodic Phonology. Phonology* 5, pp. 161-168.

Johansson, S. (1973) Linking and intrusive /r/ in English: a case for a more concrete phonology. *Studia Linguistica* 27, pp. 53-68.

Jones, D. (1972) *An Outline of English Phonetics.* Cambridge: CUP.

Kager, R. (1993) Alternatives to the iambic-trochaic law. *Natural Language and Linguistic Theory 11,* pp. 381-432. (in Kager 1999)

Kager, R. (1995) Review of Hayes 1995. *Phonology 12,* pp. 437-64. (in Kager 1999)

Kager, R. (1999) *Optimality Theory.* Cambridge: CUP.

Kager, R. (2000) Stem stress and peak correspondence in Dutch. In J. Dekkers et al. (eds.) *Optimality Theory: Phonology, Syntax and Acquisition.* Oxford: OUP, pp. 121-150.

Kahn, D. (1980) *Syllable-Based Generalisations in English Phonology.* New York: Garland.

Kaisse, E.M. (1985) *Connected Speech: the Interaction of Syntax and Phonology.* Orlando: Academic Press.

Kaye, J. and J. Lowenstamm (1984) De la syllabicité. In F. Dell et al. (eds.) *Forme Sonore du Langage.* Paris: Hermann, pp. 123-159.

Kenstowicz, M. (1970) On the notation of vowel length in Lithuanian. *Papers in Linguistics 3,* pp. 173-214. (in Donegan 1985)

Kenstowicz, M. (1981) A note on cyclic stress in Levantine Arabic. In H. Borer and Y. Aoun (eds.) *Theoretical Issues in the Grammar of Semitic Languages.* Cambridge, Massachusetts: MIT Department of Linguistics, MIT Working Papers in Linguistics (vol. 3), pp. 21-31.

Kenstowicz, M. (1983) Parametric variation and accent in the Arabic dialects. In A. Chukerman, M. Marks and J.F. Richardson (eds.) *CLS 19.* Chicago: Chicago Linguistic Society, pp. 205-213.

Kenstowicz, M. (1986) Notes on syllable structure in three Arabic dialects. *Revue Québecqoise de Linguistique 16,* pp. 101-128.

Kenstowicz, M. (1994) *Phonology in Generative Grammar.* Cambridge MA/ Oxford UK:Blackwell.

Kenstowicz, M. and C. Kisseberth (1977) *Topics in Phonological Theory.* New York: Academic Press.

Kenstowicz, M. and C. Kisseberth (1979) *Generative Phonology: Description and Theory.* New York: Academic Press.

Kiparsky, P. (1979) Metrical structure assignment is cyclic. *Linguistic Inquiry 10,* pp. 421-442.

Kiparsky, P. (1982a) From cyclic phonology to Lexical Phonology. In H. van der Hulst and N. Smith (eds.) *The structure of Phonological Representations (Part II).* Dordrecht: Foris, pp. 131-76.

Kiparsky, P. (1982b) Lexical morphology and phonology. In I.S.Yang (ed.) *Linguistics in the Morning Calm,* pp. 3-90.

Kiparsky, P. (2000) Opacity and cyclicity. *The Linguistic Review 17,* pp. 351-365.

Kisseberth, C. (1970) On the functional unity of phonological rules. *Linguistic Inquiry 1,* pp. 291-306.

Kouloughli, D. (1978) 'Contribution à la Phonologie Générative de l'Arabe: Le Système Verbal du Parler Arabe du Sra'. Thesis of the third cycle, Paris VII. (in Angoujard 1990)

Kreidler, C.W. (1989) *The Pronunciation of English.* Oxford: Blackwell.

Kuroda, S.-Y. (1967) *Yawelmani Phonology.* Cambridge, Mass.: MIT Press. (in Noske 1993)

Ladefoged, P. (1962) *Elements of Acoustic Phonetics.* Chicago: University of Chicago Press. (in Kreidler 1989)

Ladefoged, P. and I. Maddieson (1996) *The Sounds of the World's Languages.* Oxford:Blackwell.

Lapointe, S.G. and M.H. Feinstein (1982) The role of vowel deletion and epenthesis in the assignment of syllable structure. In H. van der Hulst and N. Smith (eds.) *The Structure of Phonological Representations (Part II).* Dordrecht: Foris.

Lass, R. (1976) *English Phonology and Phonological Theory: Synchronic and Diachronic Studies.* Cambridge: CUP.

Laufer, A. (1991) The "glottal fricative". *Journal of the International Phonetic Association 21,* pp. 91-3. (in Ladefoged and Maddieson 1996)

Lehiste, I. (1970) *Suprasegmentals.* Cambridge, Mass.: MIT Press. (in Ladefoged and Maddieson 1996)

Lenneberg, E.H. (1970) The biological foundations of language. In M. Lester (ed.) *Readings in Applied Transformational Grammar.* New York: Holt, Rinehart and Winston.

Lepschy, A.L. and G. Lepschy (1977) *The Italian Language Today.* London: Hutchinson.

Lerdahl, F. and R. Jackendoff (1983) *A Generative Theory of Tonal Music.*Cambridge, Mass.: MIT Press.

Levin (1985) 'A Metrical Theory of Syllabicity'. Ph.D. dissertation, MIT, Cambridge, Massachusetts.

Liberman, M. (1975) 'The Intonational System of English'. PhD dissertation, MIT, Cambridge, Massachusetts. Distributed by Indiana University Linguistics Club.

Liberman, M. and A. Prince (1977) On stress and linguistic rhythm. *Linguistic Inquiry* 8, pp. 249-336.

Lyche, C. (1979) Glides in French: questions for natural generative phonology. *Lingua* 49, pp. 315-330.

McCarthy, J.J. (1979a) 'Formal Problems in Semitic Phonology and Morphology'. Ph.D. dissertation, MIT, Cambridge, Massachusetts.

McCarthy, J.J. (1979b) On stress and syllabification. *Linguistic Inquiry X,* pp. 443-466.

McCarthy, J.J. (1981) A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry XII,* pp. 373-418.

Mc Carthy, J.J. (1988) Feature geometry and dependency: a review. *Phonetica 43,* pp. 84-108.

McCarthy, J.J. (1995) Faithfulness and reduplicative identity. In J. Beckman, L. Walsh Dickey and S. Urbanczyk (eds.) *Papers in Optimality Theory.* University of Massachusetts Occasional Papers in Linguistics 18. Amherst, Mass.: Graduate Linguistic Student Association, pp. 249-384.

McCarthy, J.J. (1998) Sympathy and phonological opacity. Ms., University of Massachusetts, Amherst.

McCarthy, J.J. and A. Prince (1986) Prosodic morphology. Ms., University of Massachusetts, Amherst and Brandeis University, Waltham, Mass.

McCarthy, J.J. and A. Prince (1990) Foot and word in prosodic morphology: the Arabic broken plural. *Natural Language and Linguistic Theory 8,* pp. 209-283.

McCarthy, J.J. and A. Prince (1993) Generalized alignment. In G.E. Booij and J. van Marle (eds.) *Yearbook of Morphology 1993.* Dordrecht: Kluwer, pp. 79-153.

McCarthy, J.J. and A. Prince (1995) Faithfulness and reduplicative identity. In J. Beckman, L. Walsh Dickey and S. Urbanczyk (eds.) *Papers in Optimality Theory.* University of Massachusetts Occasional Papers in Linguistics 18. Amherst, MA: Graduate Linguistic Student Association, pp. 249-384.

Maddieson, I. (1984) *Patterns of Sounds.* Cambridge: CUP.

Maddieson, I. and K. Emmorey (1985) Relationship between semivowels and vowels: cross-linguistic investigations of acoustic difference and coarticulation. *Phonetica* 42/4, pp. 163-174.

Maiden, M. (1995) Evidence from the Italian dialects for the internal structure of prosodic domains. In J.C. Smith and M. Maiden (eds.) *Linguistic Theory and the Romance Languages*. Amsterdam: John Benjamins, pp. 115-131.

Majdi, B. and M.Winston (1994) Rules of phonology: pre- or post-syllable structure?. In M.Eid, V.Cantarino and K.Walters (eds.) *Perspectives on Arabic Linguistics VI.* Amsterdam/Philadelphia: John Benjamins, pp. 185-214.

Marotta, G. (1987) Dittongo e iato in italiano: una difficile discriminazione. *Annali della Scuola Normale Superiore di Pisa* 17/3, pp. 847-887.

Mascaró, J. (1976) 'Catalan Phonology and the Phonological Cycle'. Ph.D. dissertation, MIT, Cambridge, Massachusetts.

Mester, A. (1994) The quantitative trochee in Latin. *Natural Language and Linguistic Theory 12,* 1-62. (in Prieto i Vives 2000)

Mester, A. and J. Padgett (1994) Directional syllabification in Generalized Alignment. *Phonology at Santa Cruz 3*, pp. 79-85.

Mioni, A.M. (1993) Fonetica e fonologia. In A.A.Sobrero (ed.) *Introduzione all'italiano contemporaneo: le strutture.* Bari: Laterza.

Mohanan, K.P. (1982) 'Lexical Phonology'. Ph.D. dissertation, MIT, Cambridge, Massachusetts.

Mohanan, K.P. (1986) *The Theory of Lexical Phonology.* Dordrecht: Reidel.

Montreuil, J.-P. (1983) Vowel epenthesis in romance: a metrical analysis. *Current Issues in Linguistic Theory 36*, pp. 233-244.

Morin, Y.C. (1987) De quelques propriétés de l'épenthèse consonantique. *Canadian Journal of Linguistics 32/4,* pp. 365-375.

Muljačić, Z. (1972) *Fonologia della Lingua Italiana.* Bologna: Il Mulino.

Myers, S. (1987) Vowel shortening in English. *Natural Language and Linguistic Theory 5*, pp. 485-518.

Nespor, M. (1990) Vowel deletion in Italian: the organisation of the phonological component. *The Linguistic Review* 7/4, pp. 375-398.

Nespor, M. (1993) *Fonologia.* Bologna: Il Mulino.

Nespor. M. and M. Scorretti (1985) Empty elements and phonological form. In J. Guéron and J.Y. Pollock (eds.) *Grammatical Representation.* Dordrecht: Foris.

Nespor, M. and I. Vogel (1982) Prosodic domains of external sandhi rules. In H. van der Hulst and N.Smith (eds.) *The Structure of Phonological Representations (Part I).* Dordrecht: Foris, pp. 225-55.

Nespor, M. and I. Vogel (1986) *Prosodic Phonology.* Dordrecht: Foris.

Nespor, M. and I. Vogel (1989) On clashes and lapses. *Phonology 6,* pp. 69-116.

Noske, R. (1993) *A Theory of Syllabification and Segmental Alternation.* Tübingen:Niemeyer.

O'Connor, J.D. (1971) *Better English Pronunciation.* Cambridge: CUP.

Orešnik, J. (1972) On the epenthesis rule in modern Icelandic. *Arkiv för Nordisk Filologi 87,* pp. 1-32. (in Kenstowicz 1994)

O'Rourke, V. (1992) A reinterpretation concerning the article *lo. Linguistic Analysis* 22/1-2, pp. 117-129.

Padgett, J. (1991) 'Stricture in Feature Geometry'. Ph.D. dissertation, University of Massachusetts, Amherst.

Padgett, J. (1995) Feature classes. In J. Beckman, L. Walsh Dickey and S. Urbanczyk (eds.) *Papers in Optimality Theory.* University of Massachusetts Occasional Papers in Linguistics 18. Amherst, MA: Graduate Linguistic Student Association, pp. 385-420.

Payne, D.L. (1981) *The Phonology and Morphology of Axininca Campa.* Summer Institute of Publications in Linguistics 66. Arlington, University of Texas. (in Itô 1989)

Picard, M. (1987) On the general properties of consonant epenthesis. *Canadian Journal of Linguistics 32/2,* pp. 133-142.

Pike, K.L. (1943) *Phonetics: a Critical Analysis of Phonetic Theory and a Technique for the Practical Description of Sounds.* Ann Arbour:University of Michigan Press.

Prieto i Vives, P. (2000) Vowel lengthening in Milanese. In L. Repetti (ed.) *Phonological Theory and the Dialects of Italy.* Current Issues in Linguistic Theory 212. Amsterdam/ Philadelphia: John Benjamins.

Prince, A.S. (1975) 'The Phonology and Morphology of Tiberian Hebrew'. Ph.D. dissertation, MIT, Cambridge, Mass.

Prince, A.S. (1983) Relating to the grid. *Linguistic Inquiry* 14, pp. 19-100.

Prince, A.S. (1984) Phonology with tiers. In M. Aronoff and R.T. Oehrle (eds.) *Language Sound Structure: Studies in Phonology Presented to Morris Halle by his Teacher and Students.* Cambridge, MA: MIT Press, pp.234-244.

Prince, A.S. and P. Smolensky (1993) Optimality Theory: constraint interaction in generative grammar. Ms., Rutgers University, New Brunswick and University of Colorado, Boulder.

Pyle, C. (1970) West Greenlandic Eskimo and the representation of vowel length. *Papers in Linguistics 3/1,* pp. 115-146. (in Donegan 1985)

Rammuny, R.M.A. (1966) 'An Analysis of the Differences in the Prosodies of General American English and Colloquial Jordanian Arabic and their Effect on Second Language Acquisition'. Ph.D. thesis, University of Michigan.

Rea, J.A. (1979) The problem of the French semivowels. In I.Rauch and G.F.Carr (eds.) *Linguistic Method Essays in Honor of Herbert Penzl.* The Hague: Mouton Publishers, pp. 353-361.

Rischel, J. (1987) Is there just one hierarchy of prosodic categories? In W.V. Dressler, H.C. Luschützky, O.E. Pfeiffer and J.R. Rennison (eds.) *Phonologica 1984: Proceedings of the Fifth International Phonology Meeting, Eisenstadt, 25-28 July 1984.* London: Cambridge University Press, pp. 253-259.

Rizzi, L. (1979) La teoria della traccia e processi fonosintattici. *Rivista di Grammatica Generativa 4,* pp. 165-81.

Roberts-Kohno, R.R. (1994) Vowel coalescence and hiatus in Kikamba: evidence for empty consonants. Paper presented at 25th Annual Conference of African Linguistics. Rutgers University, new Brunswick, N.J. (in Ladefoged and Maddieson 1996)

Roberts-Kohno, R.R. (1999) Derivationalism in Kikamba Vowel Hiatus Phenomena. In B. Hermans and M. van Oostendorp (eds.) *The Derivational Residue in Phonological Optimality Theory.* Amsterdam: John Benjamins, pp. 269-94.

Roca, I. and H.M. Al-Ageli (1999) Optimal metrics. In M. Davenport and S.J. Hannahs (eds.) *Issues in Phonological Structure.* Amsterdam: John Benjamins.

Rubach, J. (1993a) Skeletal versus moraic representations in Slovak. *Natural Language & Linguistic Theory 11,* pp. 625-53.

Rubach, J. (1993b) *The Lexical Phonology of Slovak.* Cambridge, Mass., and Oxford: Clarendon Press. (in Harris and Kaisse 1999)

Rubach, J. (1998) A Slovak argument for the onset-rhyme distinction. *Linguistic Inquiry 29,* 168-79. (in Harris and Kaisse 1999)

Saltarelli, M. (1983) The mora unit in Italian phonology. In *Folia Linguistica XVII,* pp. 7-24. The Hague: Mouton Publishers.

Schane, S.A. (1973) [back] and [round]. In S.R. Anderson and P. Kiparsky (eds.) *A Festschrift for Morris Halle.* New York: Holt, Rinehart and Winston.

Selkirk, E.O. (1980a) Prosodic domains in phonology: Sanskrit rivisited. In M. Aronoff and M.L. Kean (eds.) *Juncture.* Saratoga, Ca.: Amna Libri, pp. 107-129.

Selkirk, E.O. (1980b) The role of prosodic categories in English word stress. *Linguistic Inquiry* 11/3, pp. 563-607.

Selkirk, E.O. (1981a) On prosodic structure and its relation to syntactic structure. In T.Fretheim (ed.) *Nordic Prosody II.* Trondheim: TAPIR, pp. 111-140.

Selkirk, E. (1981b) Epenthesis and degenerate syllables in Cairene Arabic. In H. Borer and Y. Aoun *Theoretical Issues in the Grammar of Semitic Languages.* Cambridge, Massachusetts: MIT Department of Linguistics, MIT Working Papers in Linguistics (vol. 3), pp. 209-232.

Selkirk, E.O. (1984a) *Phonology and Syntax: the Relation between Sound and Structure.* Cambridge (MA): MIT Press.

Selkirk, E.O. (1984b) On the major class features and syllable theory. In M. Aronoff and R.T. Oehrle (eds.) *Language Sound Structure: Studies in Phonology Presented to Morris Halle by his Teacher and Students.* Cambridge, MA: MIT Press, pp. 107-136.

Serianni, L. (1989) *Grammatica italiana: italiano comune e lingua letteraria.* Torino: UTET.

Sluyters, W. (1990) Length and stress revisited: a metrical account of diphthongization, vowel lengthening, consonant gemination and word final vowel epenthesis in modern Italian. *Probus 2,* pp. 71-101. (in Bullock 1998)

Sobkowiak, W. (1990) Hiatus-breaking glide insertion in English and Polish. In J. Fisiak (ed.) *Further Insights into Contrastive Analysis.* Amsterdam: John Benjamins, pp. 255-275.

Spence, N.C.W. (1971) The French semi-vowels: a fresh angle? *Lingua* 27, 198-215.

Steriade, D. (1982) 'Greek Prosodies and the Nature of Syllabification'. Ph.D. dissertation, MIT, Cambridge, Mass. Published by Garland Press, New York, 1990.

Steriade, D. (1991) Moras and other slots. *Proceedings of the Formal Linguistics Conference of the Midwest 1,* pp. 254-80.

Steriade, D. (1995) Positional neutralization. Ms., University of California, Los Angeles.

Suleiman, S.M. (1985) *Jordanian Arabic between Diglossia and Bilingualism: Linguistic Analysis.* Amsterdam/ Philadelphia: John Benjamins.

Tesar, B. and P. Smolensky (1993) The learnability of Optimality Theory: an algorithm and some basic complexity results. Technical Report CU-CS-678-93, Computer Science Department, University of Colorado, Boulder.

Tranel, B. (1987) *The Sounds of French.* Cambridge: CUP. (in Harris and Kaisse 1999)

Vanelli, L. (1979) Una forma suppletiva dell'articolo e la sua fonosintassi. *Rivista di Grammatica Generativa 4,* pp. 183-206.

Vincent, N. (1987) Italian. In B. Comrie (ed.) *The Major Languages of Western Europe.* London: Routledge.

Vogel, I. (1999) Subminimal constituents in prosodic phonology. In S.J. Hannahs and M. Davenport (eds.) *Issues in Phonological Structure.* Amsterdam: John Benjamins.

Watson, J.C.E. (1989) 'Aspects of the Phonology and Verb Morphology of Three Yemeni Dialects'. Ph.D. thesis, SOAS, University of London.

Watson, J.C.E. (1999a) The syllable and syllabification in Modern Spoken Arabic (San'ani and Cairene). In H. van der Hulst and N.A. Ritter (eds.) *The Syllable: Views and Facts.* Berlin/ New York: Mouton de Gruyter.

Watson, J.C.E. (1999b) CVVC syllables in Arabic. *Languages and Linguistics 4*, pp. 1-17.

Watson, J.C.E. (1999c) The directionality of emphasis spread in Arabic. *Linguistic Inquiry 30/2*, pp. 289-300.

Watson, J.C.E. (2000) The syllable and syllabification. Ms., University of Durham.

Watson, J.C.E. (2002) *The Phonology and Morphology of Arabic.* Oxford: OUP.

Wijnen, F. (1988) Spontaneous word-fragmentations in children: evidence for the syllable as a unit in speech production. *Journal of Phonetics* 16, pp. 187-202.

Wiltshire, C. (1994) Alignment in Cairene Arabic. *West Coast Conference in Formal Linguistics 13*, pp. 138-153.

Wiltshire, C. (1998) Extending ALIGN constraints to new domains. *Linguistics 36/3*, pp. 423-467.

Younes, M.A. (1995) On vowel shortening in Palestinian Arabic. In M.Eid (ed.) *Perspectives on Arabic Linguistics VII.* Amsterdam/Philadelphia: John Benjamins, pp. 157-171.

Zawaydeh, B.A. (1997) On an optimality-theoretic account of epenthesis and syncope in Arabic dialects. In M.Eid and R.R.Ratcliffe (eds.) *Perspectives on Arabic Linguistics X.* Amsterdam/Philadelphia: John Benjamins, pp. 191-213.

Zwanenburg, W. (1966) Les phonèmes 'semi-vocaliques' du frençais moderne. *Neophilologus* 50, pp. 28-33.