Universiteit
Leiden
The Netherlands

Weightless Segments - A phonetic and phonological study concerning the metrical irrelevance of syllable onsets

Goedemans, R.W.N.

Citation

Goedemans, R. W. N. (1998, September 10). *Weightless Segments - A phonetic and phonological study concerning the metrical irrelevance of syllable onsets*. Holland Academic Graphics, Den Haag. Retrieved from https://hdl.handle.net/1887/20317

Version:	Not Applicable (or Unknown)
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/20317

Note: To cite this publication please use the final published version (if applicable).

Weightless Segments

A Phonetic and Phonological Study Concerning the Metrical Irrelevance of Syllable Onsets

Proefschrift

ter verkrijging van de graad van Doctor
aan de Rijksuniversiteit te Leiden,
op gezag van de Rector Magnificus Dr. W.A. Wagenaar,
hoogleraar in de faculteit der Sociale Wetenschappen,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 10 september 1998
te klokke 16.15 uur

door

Robertus Wilhelmus Nicolaas Goedemans

geboren te Haarlem in 1967

Contents

1 Introduction

1.1 Preliminaries

A sequence of speech sounds, call it a word or an utterance, is more than the sum of its parts. The speech sounds in a string are characterised by segmental features, but these are by no means the only features we can find in an utterance. Characteristic of the fusion of segments into longer stretches of speech is the emergence of *suprasegmental* or ‘prosodic’ features. Ladefoged (1982) mentions *stress*, *length*, *tone* and *intonation* as the principal prosodic features. These features may serve to mark certain linguistic boundaries in the speech stream (Lehiste 1960), focus the attention of the listener on certain elements in the utterance (among others, Ladd 1980), and change the meaning of words in an utterance (Bolinger 1958, 1972), to mention only a few of the linguistic phenomena in which prosody plays a role. This book is concerned with only two of the prosodic features mentioned above: stress and length. Length (or rather *duration*) will play a crucial role in the phonetic chapters of this thesis (chapters 2,3 and 4). A detailed description of stress will be given in section 1.2, along with a discussion of its place in contemporary phonology. To fully grasp the concept of stress, however, we must first introduce the *syllable*.

1.1.1 The syllable: a brief introduction

The syllable is traditionally considered the smallest string of segments that can serve as a prosodic unit.[1] Intuitively everybody knows what a syllable is, and anyone can count the number of syllables in any given word (save a number of notorious examples, like *predatory*, in which speaker variation causes variability in such syllable counts, see Ladefoged

---

[1] Rhymes and moras, which will be introduced later, are even smaller prosodic units. Rhymes will be argued later to replace the syllable as the prosodic unit that is the domain in certain phonological processes, and moras do not dominate strings of segments but rather single segments. However, see van Heuven (1994) on the possibility of single segments acting as prosodic domains in general.

1982:220). The intuitive importance of this conglomerate of speech sounds is evidenced by the fact that many of the writing systems that were developed in the past are syllabic, like Hebrew. The decomposition of syllables into separately written segments by the Greeks, which stands at the base of all presently used alphabetic writing systems, was far from trivial.

Perhaps surprisingly, a clear-cut definition of this intuitively attractive prosodic unit cannot easily be found. Around 200 BC, Dionysius Thrax defined the syllable as a combination of a vowel with one or more consonants.[2] This approach to the syllable is clearly an oversimplification, yet it prevailed until some more phonetically oriented definitions were put forward in the first half of the Twentieth Century. Sweet (1902), for instance, defined the syllable as follows:

"[T]he ear learns to divide a breath-group into groups of vowels (or vowel equivalents), each flanked by consonants (or consonant-equivalents)- or in other words, into syllable-formers or *syllabics* and *non-syllabics*, each of these groups constituting a *syllable*." Sweet (p. 65).

More or less simultaneously, the relationship between syllables and *sonority* was discovered. According to Jespersen (1904) the number of syllables in an utterance is equal to the number of sonority peaks. The sonority of a sound is its loudness relative to that of other sounds with the same length, stress and pitch (Ladefoged 1982:221). This definition allows us to compare the sonority of a certain segment to that of another similar segment. It does not, however, allow a sonority comparison between, for instance, voiced and voiceless segments (the latter have no pitch). In phonology it is commonplace to look upon sonority, in a more abstract sense, as a measure for the "strength" of segments. Given this abstract notion of sonority, the various types of segments we find in languages can be ordered along a phonological "strength scale" (or *sonority hierarchy*), which is based on phonological processes like lenition and fortition. Thus, vowels are more sonorous than consonants; fricatives, liquids and nasals are more sonorous than stops; and voiced segments are more sonorous than voiceless ones. Note that Ladefoged's definition does not cover all types of comparisons that can be made within the sonority hierarchy. Hence, this hierarchy is only partly supported by phonetic evidence. Unfortunately, to our knowledge, no phonetic research has been devoted to the search for more evidence for the sonority hierarchy.

---

[2] Cf. Laziczius (1961).

It has frequently been observed that the sonority of sounds in the syllable that occur before the vocalic peak is generally rising, while the sonority of sounds after the peak is generally falling. The fact that this is no more than a tendency shows that Jespersen's (1904) definition does not cover the complete set of possible syllables. Compare the sonority curves of the words 'cry' and 'sky'. In the prevocalic part of 'cry' sonority is indeed rising, but in the prevocalic part of 'sky' it is falling, creating two sonority peaks. Yet both words are said to constitute only one syllable.

Jones (1950) defines the syllable by its prominence peak, where relative sonority, length, stress, special intonation, or a combination of these, determine the prominence of a speech sound. A word should then contain as many syllables as there are prominence peaks. In view of the fact that there is no known procedure to integrate sonority, length, stress and pitch into prominence, this definition might be phonetically more accurate than Jespersen's, but it is quite impractical.

These difficulties in defining the syllable in an impressionistic or phonetic fashion do not prevent its usage in phonology though. In the early years of Generative Phonology the need for a phonological syllable was largely ignored. Chomsky & Halle (1968), for instance, propose a stress rule for English that does not refer to the syllable as a phonological domain. In contemporary phonology, however, the syllable is identified as the stress-bearing unit (Beckman 1986; Hayes 1995). Other phonological properties that can take the syllable as their domain include tone, nasalisation and pharyngealisation.

Anderson (1969), Fudge (1969), Vennemann (1972) and Hooper (1976) were the first to recognise the importance of syllabic units in phonology. They claimed that without reference to the syllable we miss some obvious generalisations in rules that apply in the environments /_{#,C} and /{#,C}_. For example, in Dutch, the voicing distinction for obstruents neutralises in word-final position. In (1), however, we can see that voiced obstruents devoice in other positions as well.

(1)	bad	: ba[t]	*ba[d]	'bath'
	hebzucht	: he[p]zucht	*he[b]zucht	'greed'
	boodschap	: boo[t]schap	*boo[d]schap	'errand'
	hardloper	: har[t]loper	*har[d]loper	'runner'

In an SPE type of rule (Chomsky & Halle 1968) the devoicing environments have to be specified separately. If we accept the syllable as a phonological unit, however, we can specify the environment for final devoicing in one generalisation through reference to the right edge of the syllable. The rule then becomes: Devoice every syllable-final obstruent.

Other areas in which the syllable emerges are language games and morphology. In language games, speakers often switch syllables within words (e.g. French *terive* for *verite* 'truth', Lefkowitz 1987). This is hard to explain if we do not recognise the syllable as a phonological unit. Worth mentioning in this respect are also the common speech errors, like "a walt miskey" for "a malt whiskey", in which we never interchange a postvocalic consonant with a prevocalic one, or vice versa. The syllable can also play a role in morphological processes. In such processes syllables may form the target for reduplication. Sometimes syllables are "cut-off " by infixation, like in Ulwa (noun + ka 'his': súːlu → súː-ka-lu 'his dog', ásna → ás-ka-na 'his clothes', not súːl-ka-u or á-ka-sna, Bromberger & Halle 1988).[3]

We have seen that: (1) speakers are intuitively aware of the presence of syllables, (2) phonological processes take syllables as their domain of application, (3) phonological rules may refer to syllable edges, (4) syllables are used in language games, and (5) syllables form prosodic components in some morphological processes. These facts serve as evidence for the claim that the syllable is part of our representation of sound structure (Blevins 1995).

So far, we have treated the syllable as a mere cluster of speech segments. If this cluster is part of the phonology, however, we must take into account the possibility of internal organisation. The syllable may have an internal structure that facilitates its incorporation in the prosody. The design of this structure must obviously be such that rules for syllable related phonological processes can be simplified through reference to units in the design. Using the autosegmental insights in the structure of phonological representations, developed by Goldsmith (1976), Kahn (1976) postulates a separate autosegmental tier on which the syllabic units reside. In his view the syllabic structure is as in (2).

(2) σ syllabic tier

c a t segmental tier

Using this kind of structure meant that he could refer to the syllable as a unit, while in the same effort resolving the, at that time still troublesome, ambisyllabicity problem. This problem involves segments that phonologically belong to the two syllables between which they are

[3] For a discussion on Prosodic Morphology in reduplication processes see McCarthy & Prince (1986).

located. Ambisyllabicity cannot easily be accounted for in a linear model, but, when using Kahn's autosegmental syllable, we can map a single consonant onto both syllables by linking it to those syllables on the syllabic tier.

In Kahn's syllabic model all the segments that constitute the syllable have an equal status. There is evidence, however, that the segmental coherence within the syllable is not symmetrical. In the next section we review this evidence and develop the syllabic model from the one presented in (2) to the model that is currently used in most phonological theories.

1.1.2 Subsyllabic divisions

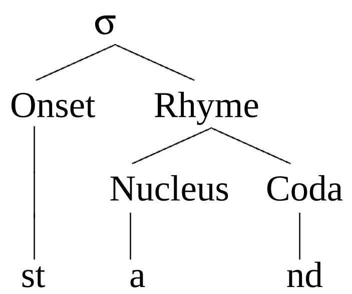
Ideas about syllable-internal constituency were first expressed by Pike & Pike (1947) who divided syllables into *margins* (sequences of consonants) and *nuclei* (vocalic sequences). This seems to be a reasonable thing to do for two reasons. Firstly, sequences of prevocalic consonants with a rising sonority can be combined to form a syllable with *any* vowel, which is taken to be evidence of a certain degree of independence between the two. Secondly, there are some co-occurrence restrictions to which such a sequence of segments has to adhere. Though rising in sonority /tl/, for instance, is not a possible prevocalic cluster in English. The existence of co-occurrence restrictions between prevocalic consonants forms evidence for their grouping into a higher order constituent. Similar considerations argue for the grouping of postvocalic margins.

Though it seems to be implied by Pike & Pike's division into nuclei and margins, we cannot look upon the postvocalic margin as a mirror image of the prevocalic margin. The co-occurrence restrictions that hold for these two clusters are, at least in English, somewhat different. The cluster /lm/, for instance, is perfectly acceptable postvocalically. Its mirror image /ml/, though, is not a possible prevocalic cluster in English (Blevins 1995). Arguments like these support the claim that prevocalic and postvocalic clusters are separate entities, and not just each other's mirror image. This claim is implicit in Hockett's (1955) division of the syllable into three, now commonly used, constituents; the *onset* (comprising the prevocalic consonants), the *peak* (or vocalic centre, nowadays mostly called *nucleus*) and the *coda* (the set of final consonants).

Contrary to what has been claimed above for the onset, the coda is not completely independent of the nucleus. We cannot combine just *any* coda with *any* nucleus. Selkirk (1978) argues that the nucleus and the coda form a higher order constituent within the syllable because there are phonotactic co-occurrence restrictions between them that do not hold between the other subsyllabic parts. A notorious example of such a restriction is the impossibility of the sequence “long vowel - velar nasal”. Combinations like [oːŋ], [iːŋ] and [aːŋ] are ill-formed in a large number of languages.

Another argument for the constituency of nucleus and coda is of a more phonetic nature. A long history of experiments shows that there is a temporal relation between a vowel and a following consonant in a large number of languages (cf. Peterson & Lehiste 1960; Chen 1970). The experiments reveal some sort of “trade-off ” relation between the nucleus and the coda, but not between the nucleus and the onset. For instance, long vowels are often followed by short consonants and short vowels by long consonants, and voiced consonants are preceded by longer vowels than voiceless consonants (cf. English *bed* vs. *bet*). These observations show that the durations of the nucleus and coda are interrelated. Following Lehiste’s (1971) assumption that such temporal relationships between two segments reflect programming as a unit at some higher level, we insert a node called the *rhyme* under the syllable node (cf. Fudge 1969; Selkirk 1978). This new node dominates the nucleus and the coda, which results in the syllabic structure presented in (3).

(3)



Not only does this rhyme unit indicate which group of segments must be identical when we create two rhyming lines of a poem, it is also very useful in many phonological rules. An example of such a rule is provided by Lass (1984). He states that, in Old English noun declensions, the onset-rhyme division is needed to account for the presence of a suffix. Let us look at some of Lass’ data.[4]

(4) a. Neuter a-stem, nom pl :

col-u	‘coals’	word	‘words’
lim-u	‘limbs’	wīf	‘women’

[4] The bars over some of the vowels in (4) indicate length.

b. Feminine o-stem, nom sg :

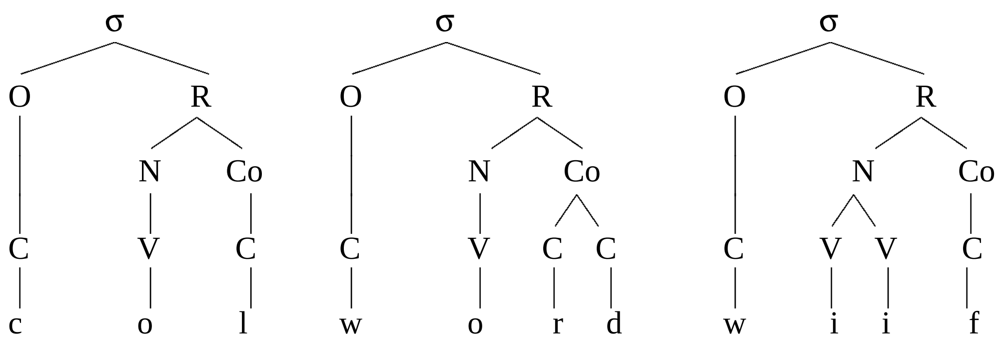
coþ-u	‘disease’	scofl	‘shovel’
far-u	‘journey’	ār	‘honour’

c. Masculine u-stem, nom sg :

breg-u	‘prince’	feld	‘field’

This pattern of Old English suffixation divides the nouns into two groups, those that end in VC and those that end in VVC or VCC. Words belonging to the former group receive a suffix while members of the latter group do not. We could create a rule for this phenomenon by simply listing the word-final syllables for which the rule applies. Such a rule, however, would not form a very satisfactory part of the phonology of Old English. With reference to the syllable, as defined by the structure in (3), a rule emerges that captures the difference between the two groups of syllables in one statement. Syllable structures of some of the words in question reveal the crucial difference between the two categories. Note that long vowels occupy two segmental slots (see Lass 1984).

(5)



It looks as if the word is somehow “weighed”, and if it is “heavy” enough no suffix is added. Heavy in this context must be taken to mean “the rhyme contains at least three segments”, which is formally expressed through the branching or non-branching of the constituents in the rhyme. The rule for Old English declension now becomes: add a suffix to words that have no branching constituent in the rhyme of the final syllable. Notice that the onset does not play a role here. The first example in (4c) shows that words with two segments in the onset can still be considered to be “light” with respect to the Old English declension rule. The considerable simplification of phonological rules that is achieved through reference to the rhyme in many other phonological processes like Old English noun declension in itself serves as evidence for such a constituent.

For most phonologists, matters such as branching constituents and weight are tightly bound to the field of stress assignment rules. It is indeed the case that (metrical) rules for stress assignment have formed the major reason to postulate the rhyme as a subsyllabic constituent in the first place. In the next section we will discuss these stress rules and their relevance to syllable constituency.

1.2 The syllable in stress rules

One of the prosodic phenomena that were mentioned in section 1.1 was stress. Just as for the syllable, we observe that stress is being recognised in phonological theory while a unified and undisputed phonetic description of the acoustic properties of stress has not yet been found. Again, the speaker's intuition on stressed and unstressed syllables in an utterance is often quite clear. Yet, these intuitions are difficult to translate into acoustic correlates that define a stressed syllable. Over the years, numerous phonetic studies have shown that there does not seem to be a single physical correlate of stress. It is more likely that a set of related correlates causes a syllable to sound stressed to a listener, and this set is probably not the same across languages. At the abstract level, however, investigations into the rules that govern the linguistic structures related to stress abound. Above, we referred to the relevance of such abstract rules for the division of the syllable into onset and rhyme. This will be discussed below, after a more elaborate introduction to stress. These considerations conclude the background that is necessary to formulate the main research question of this thesis, which will be undertaken in section 1.2.3.

1.2.1 Stress: an introduction to the phenomenon

Sweet (1902:47) defines force (or stress) by the effort with which breath is expelled from the lungs. He identifies 'loudness' as the acoustic correlate of stress. There is a, perhaps not so obvious, discrepancy between Sweet's definition of stress and his acoustic correlate. The effort with which breath is expelled is definitely speaker oriented, while loudness is a perceptually (read 'for the listener') defined property of speech that is correlated with the intensity of the speech signal.[5] This is probably what Jones (1950) had in mind when he introduced the distinction between stress (speaker activity) and prominence (effect perceived by the listener).

---

[5] But see section 6.3 for arguments that spectral slope is a better correlate of loudness.

Despite these early divisions, the focus of phonetic stress research has been primarily on its perceptual properties. An especially important example of such research is that of Fry (1955, 1958, 1965) who tried to describe stress by the perceptual strength of its acoustical correlates. In a series of related experiments Fry determined the relative strength of what he considered to be the prime acoustic correlates of stress; intensity, duration, pitch and vowel quality. He more or less systematically varied these correlates in English minimal stress pairs like *pérmit* and *permít*. The relative success of a candidate as a stress cue was determined by the percentage of listeners that judged stress to be on the syllable in which this particular correlate had been strengthened with respect to the other syllable. The first experiment revealed that, contrary to Sweet's expectations, intensity is not a good stress cue at all. It is far less effective than duration. Later experiments showed that fundamental frequency (pitch) may be an even better cue for stress than duration, and that vowel quality is the least effective of the set. Hence, Fry found the following order in the importance of stress cues: pitch >> duration >> intensity >> vowel quality.

Two of the acoustic correlates of stress that are mentioned here are very likely to be used for other purposes than stress in the phonological systems of many languages. Duration is the basis for a possible phonemic difference in vowel length, while it is also frequently used to mark the right edges of phonological phrases (Crystal & House 1988; Beckman & Edwards 1990). Pitch is the phonetic cue for tone in languages that have phonemic tonal oppositions, like Chinese. It is also the prime phonetic cue for intonation. These considerations prompt Berinstein (1979) to say that stress is **parasitic** with respect to duration and pitch; it uses the same correlates. This means for instance that, in languages that employ both stress and a phonemic length opposition, duration is less likely to be an acoustic correlate of stress, because lengthening of stressed vowels might obscure the independent vowel length contrast. An immediate consequence of this is that we recognise variability in the acoustic correlates of stress. There is no unique set of physical properties that define a stressed syllable. Rather, acoustic correlates of stress may vary in greater or lesser degree across languages (Beckman 1986; Dogil, to appear). Dogil (p.c.) suggests that the predictability of the stress position may be of some influence here. In languages that place all stresses on, say, the initial syllable, the acoustic correlates of stress may be less salient.

Pitch, or fundamental frequency, is probably the most important parasitic correlate of stress. It is not only phonemically crucial in tone languages, but it is also of significant importance in languages that use intonation as a means to encode discourse information or attitude. For instance, the fact that some part of a sentence contains relevant new information is communicated by the speaker to the listener through *focus* on that particular part (Ladd 1980). Focus is realised by an accent-lending pitch movement on the *prosodic head* of a focus domain, which can be a word or word group (Bolinger 1958; Terken 1984, 1991), as in (6a).

(6) a. What did you say?
I said [coffee]$_{+F}$
b. I wrote [tof]$_{+F}$fee, not [cof]$_{+F}$fee
c. I told you to [type]$_{+F}$ coffee, not [write]$_{+F}$ it
d. Now I heard cof[fin]$_{+F}$, not cof[fee]$_{+F}$

In (6a) the word *coffee* is in focus. The accent occurs on the first syllable, which is the prosodic head of the word. Accents also appear when we contrast two items in one sentence, as in (6b). In such cases even single syllables can be in focus. The prosodic heads that are accented through pitch movement are the stressed syllables of *toffee* and *coffee*.[6] In (6c), however, the stressed syllable of *coffee* is not in the focus domain, and therefore, not associated with a pitch change. Yet, not many speakers of English will have difficulties in identifying the stressed syllable of *coffee* in (6c). In special circumstances, it may even be the case that focus ends up on the unstressed syllable, as we can see in (6d). These examples show that the success of pitch as a correlate of stress crucially depends on the intonation of the sentence, and that pitch is not a necessary cue for stress. They also show that it is imperative that we consider stress and *accent* to be two separate linguistic phenomena (Beckman 1986; Sluijter 1995). In the remainder of this thesis we view (pitch) accent as the phonetic realisation of prominence in speech, through which the speaker conveys focus to the listener in languages like English and Dutch.

So far we have not been able to give a proper definition of stress. Considering the difficulties in pinpointing the acoustic properties of stress, we may assume that a phonetic definition of stress that holds true for all languages, and takes into consideration all the acoustic correlates, will be difficult to find. The fact that stress can be expressed in so many phonetically different ways probably means that we are dealing with an abstract phenomenon here. The observation that stress is the *relative* force of a syllable with respect to the other syllables in the word (Sweet 1902) supports the view that stress cannot be locally defined as a set of syllabic

[6] Van Heuven (1994) defends the claim that only the first segment of the word is in focus in these cases.

properties. In this light we define stress as a structural linguistic phenomenon by which the relative strength of the syllables in a word, or larger prosodic unit, is specified. The abstract phonological structure organises the syllables in a word such that the syllable that is strongest relative to the others is always the one that bears the accent when the entire word is placed in narrow focus, as in (6a).

One of the major advances that came with the introduction of *metrical phonology* was the realisation that the abstract phonological structure that determines the relative strength of the syllables in a word shows substantial formal parallels with the rhythmic structures that we can find in music and verse (Liberman 1975; Liberman & Prince 1977). As such, stress is the linguistic manifestation of rhythm.[7] The variance we find in the acoustic correlates of stress across languages can easily be covered by such an abstract notion as rhythm, since it is, by nature, a phenomenon that can be expressed in many physically different ways. The domain of metrical phonology is the formulation of the rhythmic rules that derive the possible structures defining the degrees of stress of the syllables in a word.[8] The languages of the world vary greatly in their organisations of stressed and unstressed syllables (cf. Goedemans, van der Hulst & Visch 1996). It is the task of the metrical phonologist to account for all possible stress positions, while excluding non-occurring patterns.

According to the way in which we derive their stress patterns, languages can be categorised into groups. In some languages the location of stress is, in principle, unrestricted and not predictable by rule. This type of stress may be phonemic if the language contains minimal pairs that differ only in the position of the stress, which is marked in the lexicon. Russian is the prototypical example of a language with unpredictable stress, which is often referred to as *lexical* stress. Since this type of stress can, by definition, not be derived by rules, the languages that employ it are less interesting for metrical phonologists, though the

---

[7] Postulating rhythmic alternation as the driving force for stress assignment allows for the possibility of several stresses occurring in a single word. Indeed, many of the stress languages we know show some kind of alternation between stressed and unstressed syllables on the word level. Yet, only one syllable can bear the accent when the word is in focus, which means that there must be two types of stress. We postpone the division between these two types of stress until chapter 5. In this chapter we simply refer to the position of the syllable that receives the accent when the word is in focus as the location of stress.

[8] In this section only a coarse description of some possible rules is given to serve the introduction to the main topic of this thesis. More detail on metrical rules can be found in chapter 5.

fact that lexical stresses often occur only near one of the word edges in some of these languages deserves some attention.[9]

Many languages use stress to mark the edges of words or phrases. The most simple type of rule that refers to edges is a rule that always stresses the same syllable at one of the word edges. Possibilities that are frequently employed by natural languages are: stress the first (or sometimes the second) from the left, or the first, second or third from the right. A large percentage of the languages that have such a stress system do indeed always place stress on the same location in every word, like Czech (Kučera 1961) which invariably stresses the first syllable of each word. Traditionally, the languages in the Czech-type group are said to have *fixed* stress. However, a subset of languages that marginally vary the position of stress is also (confusingly) included in this group. In such languages the position of the stressed syllable may vary (usually within a trisyllabic window at the left or right word edge), but can always be derived through application of a set of stress rules. These stress rules refer to the edges, but use secondary criteria to determine the exact location of the stress. In this set of secondary criteria we find, among other things, references to “odd or even number of syllables between stress and left or right edge”, “morphological structure” and “rhythm”, but also to “syllable weight”. The fact that the stress position can be determined through rule application is what the two language types in the fixed stress group have in common.

Languages that use syllable weight to determine the location of stress are grouped together in the set of *quantity-sensitive* languages, which has a complementary set of languages like Czech, which are *quantity-insensitive*. An example of a quantity-sensitive (QS) language is

---

[9] Revithiadou (forthcoming) notes that in many lexical stress systems the location of stress is restricted to one of the word edges, so that their stress is fixed in some sense. Notably, some truly fixed stress languages like English and Dutch do have a lexical “feel” about them, as is evidenced by minimal pairs like *súbject* and *subjéct*. The subset of words that have unpredictable stress in English and Dutch is relatively small though. In Dutch, for instance, stresses in 85% of the vocabulary can be predicted by rules (Langeweg 1988). Hence, only 15% of the words needs to be specified for stress position in the lexicon.

Idsardi (1992) proposes a formalism in which these lexical stresses are perhaps more easily dealt with. In his framework, heads and edges are parametrised instead of the foot based parametrisation we find in Haysian metrical phonology. Since lexical systems clearly have heads, Idsardi's formalism might be applied here with more success. Though the choice between head/edge or foot based analyses is a matter of some debate in contemporary metrical phonology, the issue has no bearing on this thesis, so we will leave it at these remarks (though one may note the resemblance between head/edge parametrisation and the alignment constraints in Optimality Theory that we will introduce in chapter 6).

Walangama, an Aboriginal language from Queensland, Australia (Appendix D; Tindale 1938). Consider the data in (7).

(7)	knómora	‘ear’	iŋgóːla	‘one’
	áriŋga	‘black cockatoo’	arpáːra	‘tomahawk’
	íːra	‘tooth’		

We observe that stress falls on one of the first two syllables. If one of them contains a long vowel (indicated by ː) it is stressed, if not, the first syllable carries the stress. The traditional description for stress location in languages like Walangama is something like “stress the second syllable if it is *heavy*, else stress the first syllable”.[10] Heavy syllables in this respect are the ones that contain a long vowel, *light* syllables are all the others. Syllable weight, as used in the stress systems of many other languages like Walangama, is the cue to the division of the syllable in onsets and rhymes on the basis of metrical rules. In the next section we will see how.

1.2.2 Syllable weight in quantity-sensitive stress rules

“...in all languages known to us, stress assignment rules are sensitive to the structure of the syllable rime, but disregard completely the character of the onset” Halle & Vergnaud (1980).

This quote from Halle & Vergnaud expresses the main theoretical reason for the postulation of constituency below the syllable level. Natural languages may differ in what kind of syllables they call heavy, but they all agree on the fact noted by Halle & Vergnaud. Consider (8), in which we present some possible divisions between heavy and light syllables that languages may use in their stress rules.

(8)	*Heavy*	*Light*	*Example*
	CVV	CV, CVC	Walangama
	CVV, CVC	CV	Latin

The opposition between heavy and light syllables in these examples shows variation in the vocalic and postvocalic parts of the syllable. To

[10] There are no words with long vowels in both the initial and the second syllable in Walangama (at least not in Tindale’s 1938 word list from which we inferred the stress rule). In languages that do have such words main stress may fall on the first of two heavy syllables or on the second, the stress rule must state which.

our knowledge there is no language that opposes, for instance, heavy CVC and CVV syllables to light VC and VV syllables.[11] Hence Halle & Vergnaud's claim that quantity-sensitive rules can only refer to the syllable rhyme. This restriction on the number of possible weight oppositions forms strong empirical motivation for the syllable structure presented in (3). With the help of (3) we can easily see what the structural reference to weight must be. In the same vein as in the Old English example in (5), the syllables that are potentially heavy have a branching node under the rhyme. In the second case in (8), however, branching of the rhyme itself (into nucleus and coda) can also make a syllable heavy.

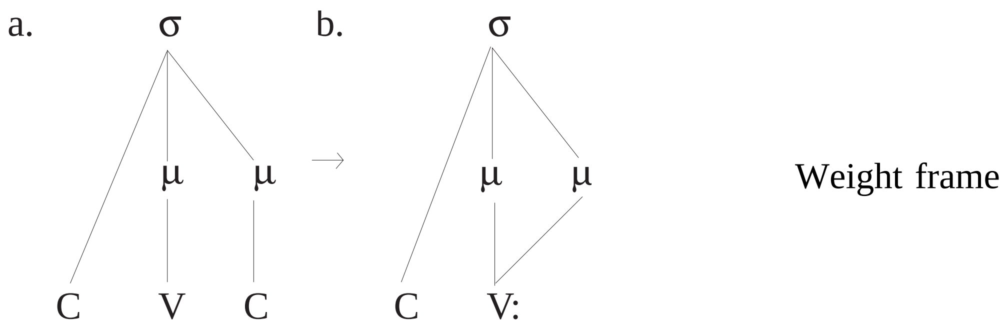
Adoption of the syllable structure in (3) has by no means been the last move concerning subsyllabic constituency in metrical phonology. Arguments for another kind of representation come from proponents of **moraic theory**, advanced by Prince (1983) and Hyman (1985), and further developed by van der Hulst (1984), McCarthy & Prince (1986), Hayes (1989), Ito (1989) and Zec (1988). Firstly, they argue that with structures like the one in (3), and similar structural descriptions, we are able to refer to the separate segments that constitute the syllable, and count them. Thus, we could count segments and evaluate /ta/ and /at/ as equally heavy. It appears, however, that processes for which such counting is needed do not feature in the phonologies of the world's languages. Usually, phonological processes that consider syllable weight act like the stress rules discussed above and ignore the onset completely. Adoption of the structure in (3) means that we must add the extra stipulation that onsets do not count.

The second argument against the structure in (3) comes from **moraic conservation**. This term covers processes that delete or shorten coda segments while simultaneously lengthening the nucleus, which can easily be found in the world's languages. Processes that delete the onset and consequently lengthen the nucleus are absent. Were we to replace the tree structure portion of the representation in (3) by a separate weight frame, on which the onset has no structural position, we could easily describe these processes as a change in the alignment of the segments to the weight frame, as in (9). The basic idea of this approach is simple: Represent segments that are prosodically active on a separate, moraic, weight tier by a unit of measure, called *mora* (μ). These moras are, in their turn, linked to the syllable. Deletion of, for example, a coda segment does not necessarily lead to deletion of its durational slot, which may lead to association of the other rhyme segment to this slot, and hence,

---

[11] But see chapter 5 for a discussion of some apparent counterexamples to this claim.

lengthening of that segment. Notice that deletion of an onset would not trigger the effect.[12] If we delete the onset, no moraic position remains for the vowel to link onto and thus lengthen. And indeed, as we noted above, no such process occurs in natural languages.

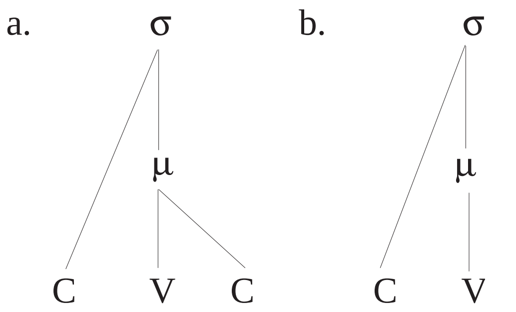
(9) 

When bound by structures like the one presented in (3), however, we can only refer to the fact that the coda triggers the effect, while the onset does not, as accidental.

The stipulation that onsets do not count in weight sensitive phonological processes is incorporated in the structure in (9). Only elements that actually count in such phonological processes may be represented on the weight frame. This is the main difference between (3) and (9). The metrical structures that determine the placement of stress are built on the moras, and phonological processes that refer to weight can only count moras. Other possible representations of syllabic structure in such a moraic framework could look like (10), where both syllables are light (1 μ) while both structures in (9) represented heavy syllables (2 μs).

[12] Much more can be said about the moraic structures in (9). Originally, Hyman (1985) linked the onset to the first mora instead of to the syllable node. Intuitively, we feel that the onset can then at least indirectly receive weight. We will show in this thesis that the onset is indeed also phonetically weightless and that we do not need onset weight in phonology. We therefore abide by the structure in (9), which is also defended by Hayes (1995) on phonological grounds.

Van der Hulst (1984) separates length from weight and inserts an extra layer of slots between the moras and the segments. On this layer the onset also has a slot. This move is defended by Lahiri & Koreman (1988) who claim that this separation is needed to describe several phonological phenomena. Dutch long vowels, for instance, need two length slots, but since they are considered light for stress purposes, they are dominated by only one mora. In the remainder of this thesis we will use the representations in (9). That does not signify a choice between these two possible representations, though. It is merely the case that we will not need a separation between length and weight, so we can use (9) for ease of exposition. In chapter 5 we will come back to the issue of moraic representations.

(10)


As we have seen above, some languages have heavy CVC syllables, while others do not. The difference between the syllabic structures of these languages is visualised by the difference between (9a) and (10a). In their basic form, all syllables have the latter structure, but, if codas need to have weight, we invoke the *Weight-by-Position* rule (Hayes 1989). This rule merely assigns a mora to the coda, realising its own representation on the weight tier. We thus get the structure in (9a).

Notice that deletion of an onset in either (9) or (10) means that any reference to the position it occupied is impossible. There seem to be phonological processes though, that do refer to this position (without considering its weight)[13]. For other phonological processes, adoption of the onset-rhyme structure in (3) as opposed to the moraic structure is claimed to be crucial (eg. Rubach, to appear). Moraic inconsistencies, such as the ability of codas to add weight in one phonological process of a certain language while in another process in the same language it is necessarily weightless, or the difference in weight between sonorant and non-sonorant codas that we find in the stress systems of languages like Inga (Levinsohn 1976), present a further challenge to moraic theory (Broselow 1995). Considerations like these have led van der Hulst & Rowicka (1997) to the claim that proponents of onset-rhyme syllables and those defending moraic syllables are referring to different phonological domains. They argue that both syllable types exist. The onset-rhyme syllable is used in lexical rules while the moraic syllable is typical for postlexical phonology. The moraic structure, then, is the one that we use when we refer to the syllable in prosodic rules. In the remainder of this thesis we will assume that the prosodic syllable is indeed moraic. In chapter 4 we will return to the postlexical moraic syllable and try to determine its exact role in metrical phonology.

[13] See Davis (1995) for a discussion of some of these processes.

1.2.3 A question about weightlessness

In the previous sections several phonological rules have been discussed that use a weighting process to determine whether a certain syllable is a valid candidate for the application of that rule. The fact that these weighting processes, together with all other known cases of phonological weighting, ignore the presence of the onset, served as evidence for subsyllabic constituency that separates the onset from the rest of the syllable, either through the postulation of an onset and a rhyme part, or by depriving the onset of a position on the moraic (weighting) tier. This split below the syllable level is certainly correct at the observational level; the empirical evidence for it is overwhelming. To our knowledge, however, a real explanation for it has never been found. The search for such an explanation is what is behind the central question we address in this thesis: *Why is the syllable onset weightless?*

Since many phonological observations find their explanation in phonetics, that is an obvious domain in which one can look for the answer. Phonetic experiments might show that there is an acoustical difference between the onset and the rhyme that serves as an explanation for the observed difference in their possible contribution to phonological weight. As we have noted above, the proposed unit for phonological weight is the mora. Though the relation between actual phonological length and the mora is disputed by some (cf. Perlmutter 1995) we take the mora to be at least an indicator of quantity (weight) or length. The acoustic correlate of length is obviously duration. Since nuclei and codas can have moras while onsets cannot, we expect some sort of difference in the durations of onsets as opposed to nuclei and codas to emerge when we compare them in phonetic duration experiments. The first half of this thesis describes some of these experiments. The next section introduces these experiments in more detail. Finally, a phonological problem is introduced that is intimately related to the onset weightlessness hypothesis, namely that of languages that do seem to have onset weight. It is this problem that we will discuss in full detail in the second half of this thesis.

1.3 Preview

1.3.1 The phonetics: production and perception experiments

Possible phonetic explanations for the weightlessness of the syllable onset can roughly be divided into two types. These types are related to the speaker and listener oriented definitions of stress that were given above. The first type is centred around some regularities we find in the durational patterns of the segments in the speech wave as it is produced by the speaker. A general tendency we find for units in a certain utterance is that their duration shortens as the number of such units present in a higher order prosodic constituent increases. In short, the more segments one puts in a syllable the shorter each will be (Nooteboom 1972; de Rooij 1979). If we take this tendency to the extreme, we might postulate the existence of prosodic units in which such shortening is so drastic that an increase in daughter units leads to a proportional shortening of all other daughters (such that the total duration of the higher-order unit does not change). An example of this would be a language in which monosyllabic, disyllabic and trisyllabic (etc.) words are equal in duration (syllables being the daughters, the word the higher order unit). Such higher order prosodic units are durationally *invariant*. It is expected that these durationally invariant units cannot exert any influence on the phonological counterpart of duration: quantity or weight (what is not there cannot serve contrastively). A first hypothesis about the weightlessness of the syllable onset could, therefore, be that the onset is durationally invariant, while the nucleus and the coda (which do show contrastive weight) are not. We would have to show, then, that the duration of onsets does not change if we increase the number of segments, while the duration of the coda significantly increases when we add coda-segments. In chapter 2, two phonetic experiments designed to test this initial hypothesis will be discussed. In doing so we mainly concentrate on the differences between the onset and the coda, disregarding the nucleus for reasons of compatibility.

The second type of explanation is perceptual in nature. Irrespective of the outcome of the first experiment, the reason for the weightlessness of the onset might be (partly) psychophysical. In the second set of experiments, described in chapter 3, we test the hypothesis that the human ear is more sensitive to duration changes in nuclei and codas than to such changes in onsets. If the explanation for the observed differences in weight is indeed psychophysical, we expect duration changes in the onset to be perceived poorly while duration changes in the nucleus and the coda should be perceived correctly (or even be exaggerated). In other words, we test the difference between segments that can, in principle, receive a mora and those that cannot, by determining for each category the perceptual saliency of the mora's phonetic correlate: duration. Listeners should be more sensitive to the duration of segments that may receive weight than to duration of necessarily weightless segments. If that is the case, and the perceptibility of its phonetic correlate is any measure for the success of a phonologically contrastive feature, then the poor perception of onset duration could be the cause of the fact that onsets cannot add phonological weight.

Considering the above, we might conclude that poor duration perception in onsets does not explain why onsets receive no mora, but rather, that the poor perception is *caused* by the fact that onsets are not marked for weight in the abstract representation of the syllable. In that case a phonetic explanation for the weightlessness of the syllable onset will be difficult to find. If we do find an effect in chapter 3, this effect would have to be of a *general* psychophysical nature to serve as the phonetic reason for the absence of onset weight. If the effect can also be found for non-speech signals, it cannot be the case that poor duration perception in onsets is caused by a structural difference between onsets and rhymes in the abstract representation of the syllable. In chapter 4 we discuss a final perception experiment that was conducted to shed some light on this matter. In the first sections of that same chapter we describe three phonetic experiments that we set up to test an extension of the perceptual hypothesis. We believe that, if there is a difference between onsets and rhymes with regard to duration perception, this might be caused by the fact that the human ear is more sensitive to duration after a certain salient point in the syllable, or rather, any auditory stimulus that shares certain characteristics with syllables. In the first two experiments described in chapter 4 we isolate three possible candidates that could serve as this "most salient point", namely: intensity peak, p-centre and CV-transition, and test whether their location has any influence on duration perception in the syllable.

1.3.2 A phonological problem

In section 1.2.3 we claimed that evidence for the weightlessness of the syllable onset is overwhelming. Indeed, most of the QS languages we know do not refer to onset weight. However, "overwhelming" does not mean that onsets are necessarily weightless in any language we encounter. Languages might exist in which the presence of an onset can make the syllable heavy for a stress rule. In fact, Davis (1985) introduces some of these languages. Among others, he mentions Western Aranda (Australian) which stresses the first syllable that begins with a consonant. The existence of such languages forms an embarrassment for phonological theories that rigidly rule out the possibility of onset weight. If we find that onsets do not add weight for a fundamental phonetic or psychophysical reason, we expect QS stress rules to ignore the onset indeed, but maybe not exceptionlessly. Depending on the nature of the phonetic correlate of weightlessness that we may discover, the onset is "rigidly weightless" or just a far worse carrier for weight than nucleus and coda. In the former case we will have to show that these languages can be reanalysed without the need for onset weight. In the latter case the exceptional languages can just be said to employ an unlikely weight factor. However, also in these cases, a more elegant reanalysis without onset weight would serve the coherence of metrical rules in general.

Other languages introduced by Davis (1985) are Mathimathi (Australian) and Pirahã (Amazonian). In these languages the *identity* of the onset consonant can play a role in the stress rules. This is not necessarily a problem for metrical phonology, which, in any case, has to formulate a rule mechanism for languages in which the identity of the coda consonant or the nucleus influences stress rules (cf. Inga above). This mechanism could then be extended to cover onset influences. As a proposal for such a mechanism, Hayes (1995) makes a clear distinction between weight and phonological *prominence*. In his view, prominence covers the complete set of possible segmental properties that can influence stress placement. It is represented by reflections of prominent segmental properties on a separate autosegmental tier to which the stress rules can refer. Chapter 5 integrates the phonetic results of the previous chapters into a slightly modified view of Hayes' phonological prominence. After a more formal introduction to the rules that are used in metrical phonology, this notion of prominence is used to combat cases of supposed onset-sensitive stress rules. An attempt is made to show that these cases of onset-sensitive stress do not counter the claim that the onset is weightless.

If we can deal with the prominence languages, we are left with languages like Western Aranda, in which the mere presence of an onset influences stress. Though this looks very much like the *weight* that the presence of a coda can add to a syllable, one could assume that the presence of an onset rather adds *prominence* to a syllable. For the majority of languages that have a rule that is sensitive to the presence of an onset, however, postulation of prominent, or moraic, onsets may not be necessary. In the first half of chapter 6 we try to reanalyse these related languages in a quantity- and prominence-insensitive fashion altogether. The stress rule that we will try to devise there will, in our view, be at least more natural than a QS stress rule that refers to onsets.

The second half of chapter 6 is devoted to stress in Mathimathi, a notorious example of how onset prominence seems to influence stress assignment. The way in which Mathimathi onsets are prominent, however, is not in accordance with the views on prominence that will be defended in this thesis (cf. chapters 4 and 5). Fortunately, a convincing alternative for the Mathimathi stress rule, based on the insights of Gahl (1996), can be provided. A detailed analysis of the origins of this unusual stress pattern will be presented in section 6.3.

The final case that has been presented in support of onset weight is Pirahã, which divides heavy and light syllables on the basis of presence *and* identity of the onset. Because it considers the identity of the onset, the Pirahã stress rule must already be analysed with reference to prominence. We do not view the fact that this case for onset weight is a prominence system as a coincidence. One might suspect that Pirahã is a rare case in which the presence of onsets can indirectly add prominence to the syllable. However, if we do find, in the next chapters, that reference to onset weight is merely very unlikely but still possible, the Pirahã case falls out naturally.

Finally, in chapter 7 we summarise and discuss the main findings and conclusions.

2 Onset Durations in Production Experiments[1]

2.1 Introduction

As we have stated in the first chapter, one of the aims of this thesis is to find a phonetic explanation for the phonological weightlessness of the syllable onset. In section 1.3.1 we put forward the hypothesis that this weightlessness might be the result of durational invariance of the onset. This means that the total duration of the string of consonants that make up the onset does not depend on the number of consonants in the set, but remains relatively constant. The reasoning behind this explanation is that a certain segmental or suprasegmental property cannot be phonologically distinctive in a certain unit when the measurable values of this property do not vary across different instances of this unit. Thus, if the duration of onsets remains largely the same, whatever the number of consonants present in that onset, then the onset cannot contribute to the phonological counterpart of duration: weight or quantity.

Measurements of phonetic correlates usually only roughly approximate the phonological ideal. Consequently, we do not expect to find onsets to be durationally invariant in the absolute sense. When we test the hypothesis that the onset is constant in duration we have to make sure that we do not draw false conclusions from a slight increase in onset duration that we might find if we increase the number of onset consonants. Addition of an extra segment is likely to induce a slight increase in duration. What we must determine instead is whether the duration of the onset is *relatively* invariant. In this respect it is probably rewarding to compare the durations of several onset clusters of different sizes to coda clusters that match them in size and identity of the consonants. Remember that codas are potential weight-bearing units. Hence, we do not expect them to be durationally invariant. If our initial hypothesis is correct, and the weightlessness of the onset is caused by its invariance in speech production, then we predict that durations of coda clusters will significantly increase with the number of segments while

[1] The experiments reported on in this chapter have been published in Goedemans & van Heuven (1993).

such a duration increase in onset clusters, if present at all, will be less sizeable. Two experiments that were conducted to test these predictions are described in this chapter. In sections 2.3 and 2.4 the details of these experiments will be presented. First, however, we will present some background information concerning speech-related duration experiments.

2.2 Previous research on duration in speech

Durations of prosodic units and single segments depend on many internal and external factors. These factors have been subject to numerous extensive phonetic studies. It is impossible to review them all here, so we will briefly discuss some of the most relevant ones.

Early phonetic research showed that the duration of a single segment depends on the identity of that segment. Lehiste (1970, and references cited there) states that the *intrinsic duration* of the vowel /a/, for instance, is longer than the intrinsic duration of the vowel /i/ (exact differences are language specific: for English /i/ and /æ/ Peterson & Lehiste 1960 report 206 and 280 ms, respectively). An explanation for this can be found in the greater articulatory movements that are involved in the production of low versus high vowels. For consonants the picture is less clear. There are several internal factors that determine the intrinsic duration of a consonant. Lehiste (1970) mentions place and manner of articulation as the key factors. It seems logical that a trilled /r/ has a longer duration than a single flapped /t/. Lehiste notes, however, that besides these obvious cases, no clear generalisations can be made. As far as place of articulation is concerned, labials seem to be generally longer than velars and alveolars. But for the ad hoc assumption that it takes longer to fill the oral chamber with air when the closure is labial than when it is velar or alveolar, this observation remains unexplained.

In the previous chapter a second source of influence on the duration of segments was already mentioned. In section 1.1.2 the dependency between the duration of the nucleus and the voicing of the coda was put forward as evidence for the constituency of nucleus and coda on a higher level. No such relations seem to exist between onset and nucleus, but a host of other properties of postvocalic consonants can influence nucleus duration (cf. Lehiste 1970). Therefore, the identity of neighbouring segments is identified as the second important conditioning factor for the duration of single segments.

Other influences on the duration of segments are more prosodic in nature. These influences are, therefore, not limited to single segments. Rate of speech, for instance, naturally correlates with the duration of units in the utterance. Words and phrases can be uttered at higher speed. This leads to shortening of the words and concomitant reduction in segmental duration (see Caspers 1994, and references cited there). It is unlikely that one can vary the tempo of single segments and whole words or sentences independently. Likewise, it is unlikely that single segments can lengthen under stress. English vowels in stressed positions, for instance, tend to be longer than vowels in unstressed positions (cf. Fry 1958), but stress is typically assigned to an entire syllable. It is to be expected that *all* the segments in a stressed syllable lengthen. A vowel is simply the type of segment that is most susceptible to stress-induced duration increase, because vowels have the most prominent steady-state portions in the syllable, and those portions are typically lengthened (see Clements & Hertz 1996; Vollmer 1997).

In sum, the durations of lower-order units in the speech stream depend on the prosody because they are incorporated into higher-order units that are the domains of prosodic phenomena like stressing or tempo. Prosodic influences of another type are *positional* in nature. Phonological units may be lengthened or shortened depending purely on their location in a higher-order prosodic unit. It has often been noted, for instance, that syllables may lengthen when they are in word-final position. This preboundary lengthening effect can be found in many languages. Among others, Lehiste (1980) and Nooteboom & Doodeman (1980) report it for English and Dutch, respectively. In these languages, it occurs before syntactic boundaries, dividing the utterance into logical units that are easy to process for the listener.[2]

Finally, the size of a prosodic unit affects the duration of its constituents. Nooteboom (1972) shows that the duration of a stressed /a/ in Dutch ranges from 219 ms in monosyllabic words to 121 ms in tetrasyllabic words. Such size-dependent duration differences are intimately related to the problem discussed in this chapter. Nooteboom does not claim that whole Dutch words are durationally invariant, but in later work Nooteboom & Cohen (1988) present indications that speakers seem to strive towards a situation in which the duration of a cluster of onset consonants is more or less equal to the duration of a single consonant. They illustrate this with spectrograms of the words *sop* ‘suds’, *stop* ‘stop’ and *strop* ‘noose’. These spectrograms, which are reproduced in figure 1, show that the onsets of these words have about equal durations, while spectrograms of *lief* ‘nice’, *liefs* ‘something nice’ and *liefst* ‘nicest’ show a slight increase in coda duration (unfortunately exact

[2] See also Gussenhoven & Rietveld (1992) and Cambier-Langeveld (1997) for further discussion.

duration values are not given). This increase in coda duration is confirmed by Chen (1970) who finds differences in coda durations for /paik/ and /paikt/, uttered by English speakers, of 171 ms versus 230 ms.

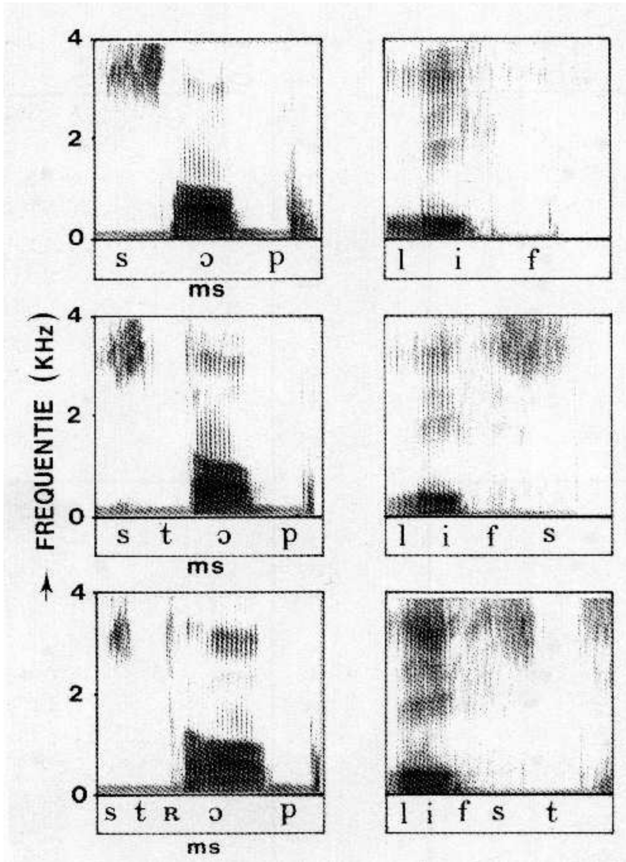


Figure 1: spectrograms of *sop, stop, strop, lief, liefs* and *liefst*. Taken from Nooteboom & Cohen (1988).

On the other hand, Lindblom, Lyberg & Holmgren (1981) show, in a series of production experiments, that the duration of the onset in a stressed second syllable of a disyllabic word depends on the number of segments in that onset. On average, the durations they find range from about 150 ms for a single /s/ through 185 ms for /st/ to 225 ms for /str/. Lindblom et al. also report shortening of segments in onset clusters with respect to monosegmental onsets (as reported by Nooteboom & Cohen for Dutch). Single /s/'s generally had a longer duration than /s/'s contained in a cluster. It seems that there is some compensation in the duration of a Swedish segment when another segment is added to the consonant cluster in which it occurs, but this effect is not nearly big enough to keep the duration of such clusters constant. Lindblom et al. conclude that *compensatory shortening* is limited. These findings are at variance with the Dutch spectrograms presented by Nooteboom & Cohen which show compensation to be rigorous.[3] The Swedish data confirm our expectation that we must not expect to find truly invariant cluster durations; according to the Lindblom et al. study the duration of onsets increases slightly with the number of consonants included in those onsets. However, no firm conclusion regarding onset weightlessness can yet be drawn from the absence of the predicted invariance. We still do not know whether the duration increase in onsets that was reported in the above studies is (phonologically) relevant. As was noted above, we can only make claims about phonological weight after we have compared duration increase in onsets to that in codas. Unfortunately, Lindblom et al. (1981) include only marginal data on codas. They present some measurements for coda durations, but do not discuss the difference between a single /s/ and a whole consonant cluster in this case. However, their data allow for some post hoc calculations. After adding up the separately listed mean durations for coda segments, we find a mean opposition of 250 ms [s] - 400 ms [rsp]. Hence, adding segments to a consonant cluster seems to have a larger durational effect on coda clusters than it has on onset clusters. These findings seem to point in the direction of a systematic weight difference between onsets and codas, but further specific studies are needed to provide solid evidence for this hypothesis.

Duanmu (1994) presents a study on Mandarin and Shanghai Chinese in which phonetic duration is explicitly used as evidence for the presence of phonological weight. He finds a steady difference of 50 ms in the average durations of Mandarin (215 ms) and Shanghai (162 ms) syllables which he uses as evidence for the claim that the former are underlyingly heavy while the latter are light. Reference to separate duration contributions of onset, nucleus and coda is not made.

To our knowledge, the first study that has been devoted to the systematic comparison of onset and coda durations for clusters with an increasing number of segments is that described in Goedemans (1993) and Goedemans & Van Heuven (1993). This study was carried out as a pilot experiment for this thesis. In the next section we will recount the procedure and the results of this experiment in detail.

---

[3] We must keep in mind that Nooteboom & Cohen (1988) report a tendency towards invariance of onset duration on the basis of three spectrograms only. It might be that these spectrograms are idiosyncratic in this respect. Verification of the trend reported by Nooteboom & Cohen for Dutch was an important incentive to conduct the first experiment presented below.

2.3 A pilot experiment

If durational invariance as a universal phonetic phenomenon is the cause of onset weightlessness, we expect it to occur in all languages. In this light it is surprising that the conclusion we can draw from the experimental results found for Swedish by Lindblom et al. (1981) diverges rather sharply from what we might conclude from the spectrograms presented by Nooteboom & Cohen (1988). It has been noted above that the true durational invariance that we find in Nooteboom & Cohen’s spectrograms is probably very difficult to reproduce, whereas the Swedish results reflect a phonetically more likely lengthening effect. Therefore, we judged an attempt at independent confirmation of the Dutch duration effect in onsets to be called for. As follows from the arguments presented above, it was also necessary to compare the results for the onset with duration values for coda clusters. The production experiment described here combined these two goals. Exactly formulated the hypotheses related to these goals are: 1. *The duration of the syllable onset remains constant, it does not depend on the number of segments in that onset* (which we will call *the Strong Hypothesis*). 2. *The duration of the syllable onset varies with the number of segments in the onset, but compared to duration variation in matching codas, the onset variation is smaller* (*the Weak Hypothesis*). If we find either hypothesis 1. **or** 2. to be true we may have found an explanation for the phonological weightlessness of the syllable onset.

The experiment was guided by the material found in Nooteboom & Cohen (1988). As we will see below, this resulted in some compromises which made it necessary to conduct a control experiment. This second experiment will be discussed in section 2.4.

2.3.1 Stimuli and method

To ensure compatibility with Nooteboom & Cohen (1988) we included in our stimulus set the Dutch words for which they present the onset spectrograms (cf. section 2.2). As an extension on their list, we added a word with an empty onset: *op* ‘on’. An empty onset is phonotactically legal in Dutch, but its position will typically be filled by a glottal stop (cf. Jongenburger & Van Heuven 1991 for discussion of this phenomenon in accented words). These glottal stops might become important if the two hypotheses presented in the previous section turn out to be false. In the light of the weight versus weightlessness discussion it might then be interesting to compare the durations of the inserted glottal stop (which can never be moraic), the segmental onsets (which are claimed to be weightless), and codas (which are claimed to be moraic). Our hypothesis would be that the absolute durations of the glottal stop and the onset are comparable, as opposed to the coda. If both the Strong and the Weak Hypothesis are falsified, then such grouping would constitute the only indication (but no more than that) of a difference between onsets and codas we could obtain from production data.

Furthermore, we included three words containing the long vowel /a/ and the /f/, /fs/ and /fst/ coda clusters that Nooteboom & Cohen use, again supplemented with a version in which the coda is empty. Finally, we repeated the coda set using a word with the short vowel /ɑ/ in the nucleus, but keeping the coda consonants the same. This was done to check whether vowel length had any influence on the duration increase in Dutch codas.

Five tokens of the following meaningful Dutch words were recorded by two native speakers of Dutch (one male, one female), in the fixed carrier sentence *Wil je* [target] *eens zeggen* /wɪl jə ... əns zɛxə/ 'Would you please say [target]' (with accent on [target]). They are presented in (1).

(1)	subset A:	op	/ɔp/	'on'
	onset, V-nucleus	sop	/sɔp/	'suds'
		stop	/stɔp/	'stop'
		strop	/strɔp/	'noose'
	subset B:	ga	/xa/	'go'
	coda, VV-nucleus	gaaf	/xaf/	'neat, unscathed'
		gaafs	/xafs/	'something neat'
		gaafst	/xafst/	'most neat'
	subset C:	laf	/lɑf/	'cowardly'
	coda, V-nucleus	lafs	/lɑfs/	'something cowardly'
		lafst	/lɑfst/	'most cowardly'

This set constitutes a non-ideal but workable compromise between the phonotactic and lexical limitations of Dutch and the full expansion of the ideal symmetrical pair of schemata $(((C_1)C_2)C_3)V_1(V_1)C_4$ and its mirror image, given the desire to deviate as little as possible from the set of target words that Nooteboom & Cohen used. We did not try to approach the ideal set through usage of nonsense words, since these were likely to affect the fluency with which the subjects could utter the sentences. In such non-fluent speech it would be difficult to determine the duration of, for instance, the glottal stop that precedes a word with an otherwise empty onset. A slight pause before this word could be misinterpreted as part of the glottal stop. Keeping in this in mind, we tried to ensure reasonable fluency by choosing two "trained" subjects who work in the phonetics department of Leiden University.

Note that, in our selection, the coda consonant and the vowel are kept constant when the onset is the target (subset A), and that the onset is kept constant (within the subset) when the nucleus and coda are the targets (subset B for coda targets with long vowels; subset C for coda targets with short vowels). This should limit the relative influence of neighbouring segments on the duration increase in the relevant clusters (cf. section 2.2). Differences in prosodic influence on duration are expected to be negligible, since the target words occupy the same accented position in the sentence in all cases.

In Dutch, empty codas are not allowed after a short vowel (hence the absence of an empty coda in subset C). An empty coda is perfectly legal after a long vowel (this results in an *open* syllable). Vowels in open syllables are not followed by a glottal stop, but they will be longer than the same vowels in closed syllables (cf. Quené 1989). Therefore, the function of the onsetless word in subset A is purely to compare glottal stop duration to segment duration in onsets and codas. Such glottal stop duration in onsets cannot be compared to glottal stop duration in codas. Hence, the relevance of the empty coda word in subset B is limited. It may serve to verify Quené's claim that vowels are indeed longer when they are not followed by a coda.

The recordings were made on a REVOX B-77 tape-recorder in a sound-attenuating recording booth with a Sennheiser MKH 416 condenser microphone. The target words were excised from the recordings and AD converted to a MicroVAX workstation (10 kHz, 12 bit, 4.5 kHz LP). After this, the segment boundaries were determined by examining the oscillograms in a high-resolution waveform editor (SESAM).[4] The boundaries of all the segments in the words were labelled, and the labelled files were stored on disk. An example of a sesam file with labels is given in figure 2. Durations of onset, nucleus and coda were measured for each word with the help of these labels.

[4] The segmentation of words is not a standard task. We followed the segmentation criteria described in Rietveld & Van Heuven (1997) as much as possible. Any effects that other, arbitrary, decisions could have on the results of this experiment should be cancelled by the consistency with which we carried out this segmentation.

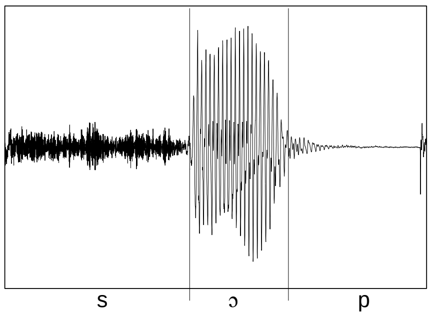


Figure 2: labelled oscillogram of the Dutch word *sop*.

2.3.2 Results and discussion

The onset and coda durations we found are plotted in figures 3 and 4, respectively, against nucleus duration (plotted horizontally) and number of segments in the target consonant cluster. The mean values for these onset and coda clusters are given in appendix A (Table I).

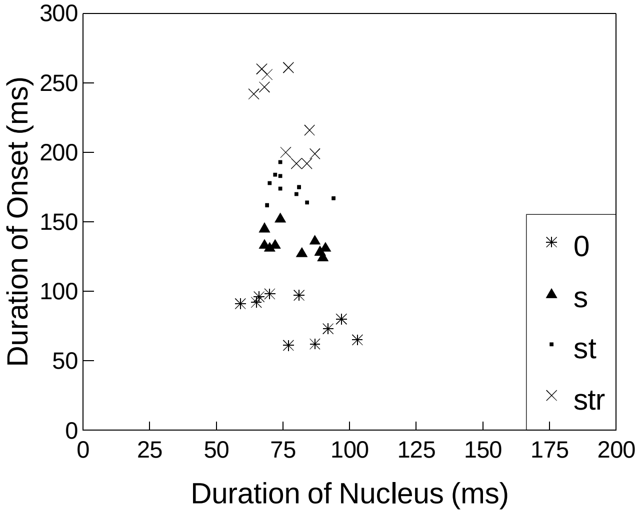


Figure 3: onset cluster duration as a function of nucleus duration and cluster size (subset A).

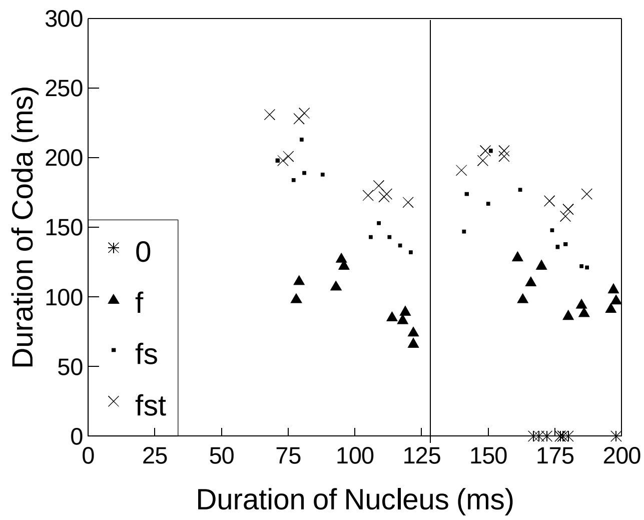


Figure 4: coda cluster duration as a function of nucleus duration and cluster size (subsets B and C, data points for long and short vowels are separated by the vertical line).

In both figures we find a clear separation of the data points for the different cluster sizes: the more segments contained in the cluster, the greater the duration of that cluster. The effect seems to be of the same order for onsets and codas alike. We also observe that, in all cases, data points indicating longer clusters are located slightly to the left of data points indicating shorter clusters. This negative correlation between nucleus duration and onset or coda duration is a measurement for the amount of shortening that the vowel undergoes if the target consonant cluster is lengthened. In figure 4 the data points for long and short vowels are easily separable (by the vertical line). For long vowels we find data points for empty codas on the horizontal axis (they have no duration). There are no data points for empty codas to the left of the vertical line since empty codas after short vowels are disallowed in Dutch.

Figure 3 shows that the duration of the onset increases with the number of segments in it. Even when empty onsets are not considered, the effect of number of segments on onset duration is significant. In a one-way analysis of variance we find $F(2,27)= 61.1$, ($p<.001$). Similarly, in figure 4 the duration of the coda increases with the number of coda consonants for both types of nucleus. Two one-way analyses of variance show the effects to be significant, $F(2,27)=38.0$, ($p<.001$) and $F(2,27)=39.4$

($p<.001$) for codas after long and short nuclei, respectively (excluding empty codas[5]). All this means that, according to these data, onset duration is *not* invariant, neither in the absolute nor in the relative sense of the term. On the basis of this experiment we would have to reject both the Strong and the Weak Hypothesis stated at the beginning of this section.

As was observed, longer consonant cluster duration is compensated for by shortening of the nucleus for both onset and coda. Crucially, however, there is less compensation in the vowel for longer onset duration (correlation coefficient: $r=-.30$, ins.) than for longer coda duration ($r=-.63$, $p<.001$ for short vowels; $r=-.68$, $p<.001$ for long vowels, again excluding empty codas). These results lend support to the claim made in chapter 1 that nuclei and codas are more intimately related than nuclei and onsets. We stated there that (significant) durational dependencies between two units reflect constituency of these units at a higher level.

We noted in section 2.3.1 that vowels before empty codas should be longer than vowels before "filled" codas. We can see in the right half of figure 4 that this prediction is more or less reflected in our data. Mean vowel durations are slightly longer before empty codas than before other codas. However, the effect does not seem to be bigger than what we might expect on the basis of the negative correlation between number of coda segments and vowel duration. We believe that Quené's (1989) observation on the longer duration of vowels before empty codas reflects the logical extreme of this correlation. Note also that vowel length does not seem to have an effect on the duration increase in the coda. Duration values by coda size after long and short vowels are of the same magnitude.

Finally, with respect to empty onset clusters, it may be noted that the durational behaviour of the glottal stop in the empty onset escapes the alternative hypothesis presented in section 2.3.1. It is not comparable in duration to any of the remaining clusters (see appendix A, table I). Even when we take into account the fact that the onset and coda segments we used may have different intrinsic durations (which is not nessecary in our case since the single segment onset and coda, /s/ and /f/ that are to be compared to the glottal stop, are similar in this respect: cf. also footnote 6), we cannot conclude that absolute onset duration is closer to the

[5] Empty onsets and codas are left out of the statistics here and below to allow us to compare the results for onsets and codas after long vowels to those for codas after short vowels (for which there is no empty variant). Furthermore, if we include the 0 ms codas after long vowels in the statistics, that will have a predictable positive influence on the clustersize by duration effect. Finally, we introduced the empty onsets and codas in the data set to test an alternative hypothesis in case our initial hypotheses proved incorrect. It seems illogical, therefore, to include them in the statistics beforehand.

duration of the glottal stop than is absolute coda duration. Hence, even the slight indication of a difference in weight between onsets and codas that we hoped to find in these production data cannot be confirmed. We attribute the difference between the empty onset duration and the “filled” onset and coda durations to the small intrinsic duration of a glottal stop compared to that of /s/ and /f/.

In conclusion, the effect of number of consonants on the duration of the syllable as a whole, as well as on the duration of the target cluster, is even larger for onsets than for codas. These results run counter to the suggestion made above that the duration of onset clusters is constant, and does not contribute to syllable duration. It seems that we cannot explain the weightlessness of the onset with evidence from production data. However, there is one possibility that we have not yet considered. In spite of what is generally assumed, some linguists, like Lindblom et al. (1981) and Davis (1985, and references cited there), claim that compensatory dependencies between onset and nucleus do exist. It is evident from figure 3 that no drastic compensation between onset and nucleus takes place in our data. Yet, claims to the contrary must be taken seriously. Before we continue we must determine if more intricate dependencies between onset and nucleus exist which might even explain the weightlessness of the onset. Below we present the results of a reanalysis of the data we obtained in the experiment discussed above. In this reanalysis we viewed the problem from a slightly different angle.

2.3.3 Ascent and Descent durations

It is to be expected that the lack of a large compensatory effect between onset and nucleus in our data only constitutes a confirmation of the general claim that no such effects exist (cf. section 2.2). However, in the preceding discussion we have overlooked one serious possibility. It might be the case that the compensatory effect holds between the onset *and only a part of the nucleus*. As we have seen in chapter 1, the syllable consists of an intensity peak flanked by two intensity minima. The intensity peak is usually located in the first half of the vowel. Suppose that this peak forms a boundary and that the vocalic part before the boundary is related to the onset while the vocalic part after the boundary is related to the coda. By far the larger part of the vowel occurs after the boundary. It might be expected, therefore, that some significant compensatory effects occur in the part of the syllable after the boundary, which will henceforth be labelled *descent*. This does not exclude a compensatory effect in the part of the syllable before the boundary, henceforth labelled *ascent*. If this compensatory effect were of a different order than the effect we find for the descent, the division of the syllable into ascent and descent finds immediate support. In the ideal case, that compensatory effect is so large that the total duration of the ascent is invariant, which would neatly explain the weightlessness of the onset. Weight could then be attributed to post-boundary constituents (including the vowel). If this hypothesis is correct, we expect some compensatory effects to occur in the descent, but the general trend should be that the total duration of the descent increases with the number of segments in the coda, while the total duration of the ascent remains constant, irrespective of the number of segments in the onset.

To test this hypothesis we reprocessed the data we obtained in the experiment described above. In each label-file we located the intensity peak and placed an additional label on that position. Ascent and descent durations were measured with the help of this label. In figure 5, ascent duration is plotted as a function of descent duration and number of segments in the onset, and in figure 6 descent durations for stimuli with long (right) and short (left) vowels are given as a function of ascent duration and number of segments in the coda. A table with the mean durations can be found in appendix A (Table II).

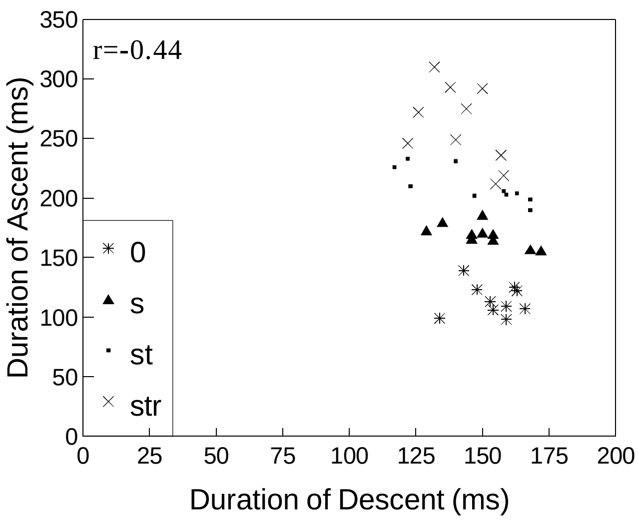


Figure 5: ascent duration as a function of descent duration and onset size

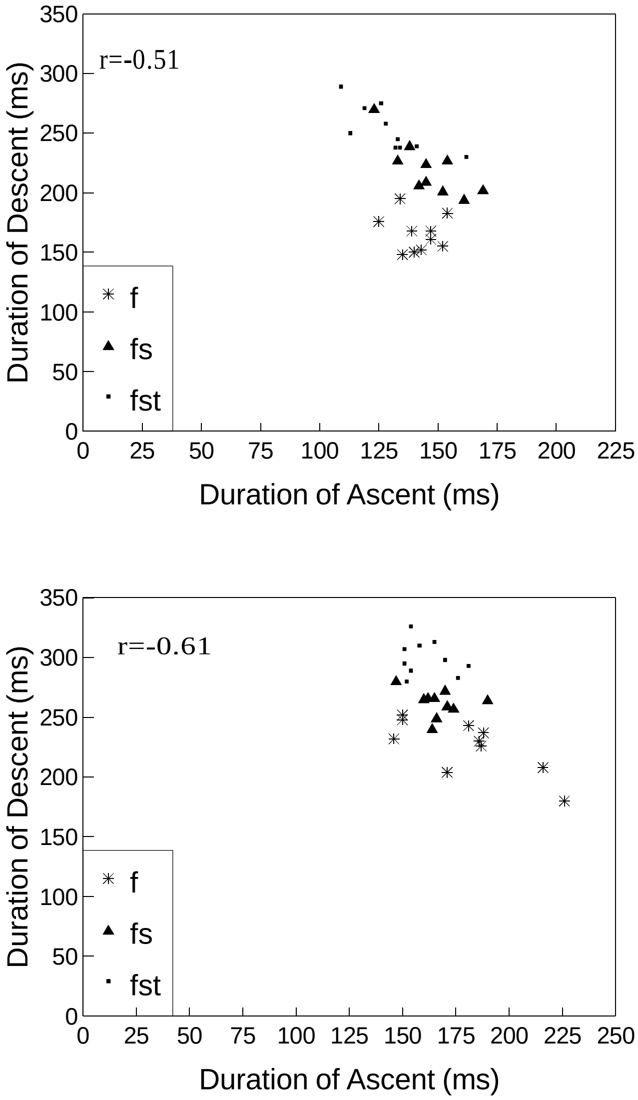


Figure 6: descent duration as a function of ascent duration and coda size for short vowels (top panel) and long vowels (bottom panel).

On a par with figure 3 we find in figure 5 that durations of ascents increase as the number of segments in the onset grows. The data points for the different onsets are about as clearly separated as in figure 3. The data presented in the two panels in figure 6 correspond to the two data sets separated by the vertical line in figure 4. Just as we found for the coda, the descent increases in duration if segments are added to the final consonant cluster for both the long and the short vowel group.

We also observe in figure 5 that lengthening of the ascent has a small shortening effect on the descent. This effect is comparable to the shortening effect that long onsets have on nuclei (cf. figure 3). Larger shortening effects can be found in figure 6. In these cases lengthening of the descent leads to shortening of the ascent.

The results presented here are fully compatible with those presented in figures 3 and 4. Like onset duration, ascent duration is not invariant. The increase with number of segments in the onset for both onset and ascent duration is as large as (or even larger than) the related duration increase in coda and descent duration, respectively. In figures 5 and 6 we thus find a confirmation of the results we found in the previous section. Ascent durations increase with the number of segments in the onset. A one-way analysis of variance reveals a significant effect of onset size on ascent duration: $F(2,27)=44.7$ ($p<.001$). Descent durations increase significantly with the number of segments in the coda: one-way effect of coda size on descent duration is $F(2,27)=52.1$ ($p<.001$) for short vowels and $F(2,27)=47.9$ ($p<.001$) for long vowels.

As in figures 3 and 4, correlation coefficients may reveal compensatory tendencies *between* the subsyllabic constituents. The strength of the shortening effect that lengthening of either ascent or descent has on the other can be expressed by their correlation coefficient. By pure observation we might conclude that such tendencies are biggest for the descent containing a long vowel. The figures confirm this: ascent $r=-.44$, $p<.01$, descent short $r=-.51$, $p<.01$, descent long $r=-.61$, $p<.001$. Thus, adding a segment to the onset results in shortening of the descent, but adding a segment to the coda results in a more drastic shortening of the ascent. We might assume that the relative resistance to shortening we find for the descent (most of the nucleus and the whole coda) is an indication of potential weight, but it may also be the case that vowels in general resist shortening more successfully than consonants. In that case the difference in the durations of the vocalic parts in ascents and descents causes the difference observed above.

Again we must conclude that onset durations are not invariant. We have determined that onsets do lengthen when the number of segments is increased. Even when we extend the domain of the onset to that portion of the syllable that may be tentatively called the “phonetic onset” (the ascent), we do not find durational invariance. However, one might object that, in our wish to closely replicate the Nooteboom & Cohen (1988) data, we have made too many concessions to the internal consistency of the stimuli. In the next section we review some possible points of criticism on this pilot experiment and present a control experiment that we carried out later with stimuli that are, in our view, not subject to such criticism.

2.4 Duration variation in mirrored clusters

In the pilot experiment described above we compared durations of onsets and codas that were perhaps not very well suited for such comparison. It is true that, when comparing onsets and codas, one tries to match two prosodically different units in any case, but in doing so, one should use units that are as much alike as possible in every other respect. Our stimuli were well suited for a comparison with the spectrograms presented by Nooteboom & Cohen (1988), but this had a negative effect on the possibilities of internal comparison between onsets and codas. Apart form their position in the syllable, the onset and coda clusters that we compared in the pilot experiment were dissimilar in two important respects.

Firstly, the segments that made up the clusters were not identical. The segments used in the onset were a fricative /s/, a plosive /t/ and a liquid /r/, while in the coda we combined /s/ and /t/ with another fricative /f/ instead of the liquid /r/. Thus, in comparing total cluster durations we ignored differences in intrinsic duration (cf. section 2.2) that might exist between /f/ and /r/. Furthermore, possible intrinsic duration differences between the monosegmental onset /s/ and the monosegmental coda /f/ were not taken into consideration. In comparing the disegmental onset cluster /st/ to the disegmental coda cluster /fs/ we ignored a possible intrinsic duration difference between /t/ and /f/.[6]

Secondly, the ordering of the segments in the cluster was not the same for onsets and codas. The plosive was the second segment to be added to the onset while it was the last segment to be added to the coda. Hence, the /t/ was cluster-medial in the maximally large onset cluster (/str/) but cluster-final in the largest coda cluster (/fst/). We do not wish to uphold that onset and coda clusters must be identical in their ordering to allow fair comparison. It is much more likely that we must observe the sonority sequencing principle (see chapter 1) which demands that segments that are closer to the vowel must be higher in sonority than those further away (though /s/'s violating this principle may occur on the syllable periphery in many languages, among others in Dutch). It is to be expected, therefore, that a well balanced set of stimuli allows only coda clusters

---

[6] Indications that some of the intrinsic duration differences mentioned here are indeed different can be found in Waals (1996) who presents durations for Dutch (onset) consonants. On average, the durations she finds are 157 ms for /f/ and 104 ms for /r/. According to her data, however, the intrinsic difference between /s/ and /f/ would be negligible. She finds a duration of 160 ms for /s/. In her data, the mean duration of /t/ is exactly equal to that of /s/. Remijsen (1996) also reports almost equal intrinsic durations for /f/ and /s/.

that are the *mirror images* of the onset clusters to which they are compared. It is self evident that the clusters in (1) do not satisfy this criterion. Albeit, true mirror images are a phonological impossibility (cf. Clements 1990), but a fair phonetic study demands the use of comparable (similar) units. The phonological prohibition mentioned above means we must compare onset and coda clusters from different monosyllabic words.

We do not know in which way these two deviations from truly comparable onset and coda clusters have affected the results of the pilot experiment. Therefore, the experiment must be repeated with stimuli that do not show these drawbacks. We must bear in mind that this exercise can only reveal *relative* invariance. Phonologically speaking, the weightlessness of the onset does not depend on the identity or the ordering of the segments in the cluster. If absolute durational invariance was the cause of onset weightlessness, we should have found it in the pilot experiment. However, the differences in identity and position of the segments in the onset and coda clusters used in the pilot experiment might well have had an adverse effect on the chances of finding relative onset invariance. The only way we can truly determine whether onset clusters have a duration that is relatively invariant with respect to coda clusters is by comparing identical clusters. Hence, the rejection of the Strong Hypothesis in section 2.3 stands. This experiment is conducted to check if our rejection of the Weak Hypothesis does not hinge on the cluster differences discussed here, and whether we may accept this hypothesis if we use comparable (mirrored) clusters. Furthermore, by using these mirrored clusters we expect to find more accurate phonetic details of the general tendencies that were revealed above.

There is one more reason, albeit of lesser importance, to conduct a follow-up experiment. This reason is related to the influence that neighbouring segments may have on the duration of segments and/or clusters. One might have noticed that the set of stimuli in (1) does not contain target onset clusters in words with a long vowel. We did not include those there because we did not expect compensatory effects between onsets and nuclei anyway (most researchers do not report it, cf. chapter 1 and section 2.2). In the light of the counter-evidence to this claim that was referred to above (Lindblom et. al. 1981; Davis 1985) the set of stimuli for the experiment described below was amended with words that had a long vowel and an onset of variable size. This allows us to determine the effect of onset duration on long vowels.

2.4.1 Stimuli and method

As in the previous experiments we used sets of Dutch words that differed only in the number of segments in the onset or coda. As was argued above, the clusters in the coda sets had to be the mirror images of the clusters in the onset-target stimuli. Though mirroring empty onsets yields no results (empty codas have no duration) we included words with empty onsets and codas anyway to allow full comparison with the pilot experiment. Base (C)V(V)(C) words were used that could be transformed into other Dutch words by just adding a segment to the onset or coda. For both onset and coda, segments were added in the order: /s/, /t/, /r/. Thus, the clusters were expanded in three steps from zero to three segments, while the other segments in the word were kept stable. For codas in words with short vowels the first (zero) stage was skipped (see section 2.3.1). To limit environmental influence on the results to a minimum the words were put in the same carrier sentence we used in the pilot experiment, in which two schwas neighboured the target word. The stimulus material is given in (2).

(2) Carrier Sentence: *wil je [target] eens zeggen* ‘would you please say [target]’

subset A:	af	/ɑf/	‘off’
onset, V	saf	/sɑf/	‘cigarette (coll.)’
	staf	/stɑf/	‘staff’
	straf	/strɑf/	‘punishment’
subset B:	aal	/al/	‘eel’
onset, VV	saal	/sal/	nonsense word[7]
	staal	/stal/	‘steel’
	straal	/stral/	‘beam’
subset C:	kwas	/kʋɑs/	nonsense word
coda, V	kwats	/kʋɑts/	‘boil (anim.), leftover (arch.)’
	kwarts	/kʋɑrts/	‘quartz’
subset D:	ma	/ma/	‘mother’
coda, VV	Maas	/mas/	‘Meuse’
	maats	/mats/	‘mates’
	maarts	/marts/	‘of March’

Unlike those in (1), the onset and coda clusters in (2) are fully symmetrical. The onsets in subsets A and B mirror the codas in subsets C and D. In the previous experiment we sacrificed symmetry to meaning, excluding nonsense words. In this experiment we wish to determine the

[7] But in some regions of The Netherlands this is the pronunciation for *zaal* ‘hall’.

durational effect for mirrored clusters, which unavoidably leads to the usage of nonsense words. The sets in (2) contain words with long and short vowels. To enable fair comparison we used the vowel /ɑ/ and its long counterpart /a/ for both onset and coda target words. Subset C has one item less because Dutch words may not end in a short vowel.

To avoid any non-fluency effects that might be caused by the usage of nonsense words we used two trained speakers who were made thoroughly familiar with the material before they commenced the experiment. The subjects were asked to read out the stimulus material twice. The digital recordings were made on a DAT tape-recorder (sampling frequency 48.1 kHz) in a small sound-attenuating recording cabin in the Phonetics Laboratory of Leiden University with a Sennheiser (MKH 416) condenser microphone.

The crucial parts of the material were edited out from the recordings after transfer to an IRIS Indigo workstation (16 kHz, 16 bits, with standard Silicon Graphics two-pole linear phase Sallen-Key low-pass filter) and stored on disk. After this, the segment boundaries were determined by examining the oscillograms in a high-resolution waveform editor (GIPOS) (cf. footnote 4). In each word, the boundaries for onsets, nuclei and codas were labelled, and the durations of the onset, nucleus and coda were measured.

2.4.2 Results and discussion

The resulting durations of the relevant target clusters are presented in figures 7 and 8 which, respectively, show onset and coda durations as a function of nucleus duration and number of segments in the cluster. Unlike figure 3, figure 7 contains data from words with a short and a long vowel (subset B). A table with the mean values for these cluster durations is given in appendix A (Table III). In both figures we find a very clear duration effect. Onsets and codas lengthen to about the same degree if they are complemented with extra segments. Like in the pilot experiment, this lengthening of onsets and codas has a shortening effect on the nucleus.

In figure 8 we can again draw a straight line to separate the data points for stimuli with long vowels from those with short vowels, though this time the line has to be tilted (compare fig. 3). The fact that the line has to be tilted is indicative of a large compensatory effect of coda lengthening on the duration of nuclei. Drawing a straight line between the measurements for long and short vowels for the onset cases in figure 7 is not possible. This is caused by the fact that the horizontal variance in the data points in figure 7 is larger than in figures 4 and 8, which might be indicative of an unrestrictive relation between nucleus duration and onset. Any onset may be followed by a vowel of any duration, while codas are more restrictive and allow less spreading in the duration of the preceding vowel. We observe that the duration increase for the left and right "groups" in figure 7 is about equal, which shows that the differences in vowel length influence onset cluster duration only minimally.

In figure 7 we can see that the duration of the onset increases in proportion to the increase in the number of segments in the cluster. As above, one-way analyses of variance were run to establish the significance of the differences. A significant effect of onset segment number on onset duration is found in this case, F(2,9)=8.0 (p=.01). In this respect the results shown in figure 7 duplicate those shown in figure 3. The extra data obtained from the long vowel-onset stimuli in subset B clearly show that onset durations also increase with the number of consonants in the cluster if the vowel is long. The effect of cluster size on onset duration for subset B is significant: F(2,9)=26.5 (p<.001).

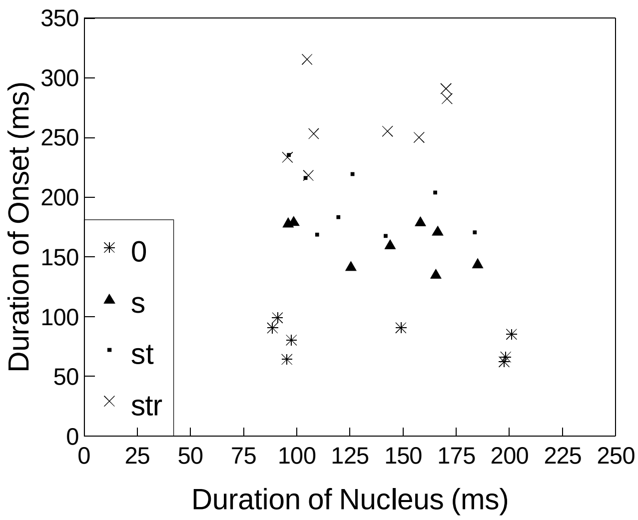


Figure 7: onset cluster duration as a function of nucleus duration and cluster size (subsets A, B).

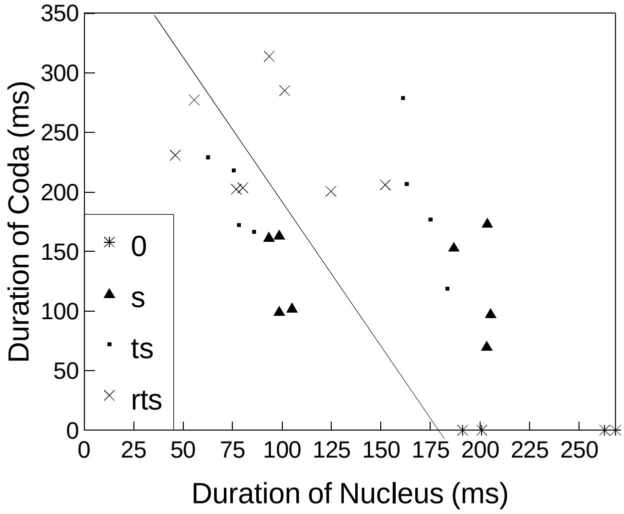


Figure 8: coda cluster duration as a function of nucleus duration and cluster size (subsets C and D, data points for long and short vowel measurements are separated by the straight line).

Figure 8 can be compared with figure 4. As far as the duration increase in the coda is concerned, the results are nearly identical. There are significant effects of the number of consonants in the coda on coda duration: $F(2,9)=8.2$ ($p<.01$) for short vowels and $F(2,9)=4.8$ ($p<.05$) for long vowels.[8] (Empty onsets and codas are again left out of the statistical calculations, cf footnote 5.) We observe that the duration increase in onsets and codas is again comparable and must conclude, therefore, that the choice of segments and the order of the segments in the cluster did not greatly influence the effects found in the pilot experiment. There is no relative durational invariance of the syllable onset, so we can safely reject our Weak Hypothesis.

A striking difference between figures 4 and 8, however, is found when we compare the size of the compensatory effect that lengthening of the coda has on the duration of the nucleus. The effect seems to be much bigger in this experiment than in the previous one, especially for long vowels. In this case the nucleus before /s/ is about 80 ms longer than the

[8] The fact that we have used an onset with two segments in subset C did not influence the results. If we compare the duration increases of the /lɑ../ words from (1C) with those of the /kʊɑ../ words from (2C) with the help of appendix A (Tables I and III), we find almost no difference.

nucleus before /rts/ whereas in figure 4 the duration difference between nuclei before /f/ and /fst/ was only about 25 ms. A similar compensatory effect cannot be found between the onset and a nucleus with a short vowel, and for the onset and the long vowelled nucleus the effect is marginal. The figures in 9 show this more clearly. There the mean duration of the nucleus is shown as a function of cluster-size and vowel length. Nucleus duration remains more or less constant for every onset type, but drops (sharply for long vowels) if we add segments to the coda. Statistics confirm what can be seen in the figures. The compensatory effect of onsets on short vowels is the reverse of what we expect; there is even a, very small, lengthening effect: correlation, $r=.18$, ins. The compensatory effect of onsets on long vowels is insignificant: $r=-.49$, ins. Compensation in the duration of short vowels before codas is quite large: $r=-.81$, ($p<.01$), for long vowels the effect is even bigger: $r=-.88$, ($p<.001$). A conclusion that can be drawn from these facts is that the choice of the segments and the position of these segments in the cluster do have an effect on the compensatory shortening of the vowel. Furthermore, these data show more clearly than the data from the pilot experiment that nuclei are durationally related to codas and not to onsets (cf. chapter 1).

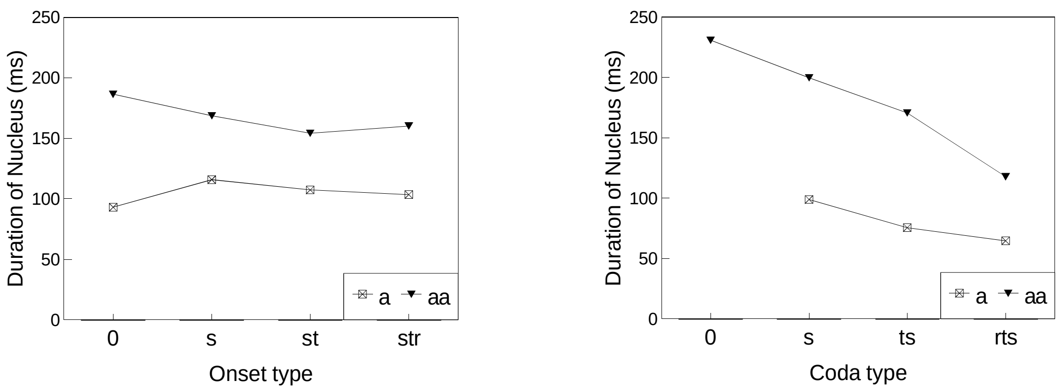


Figure 9: mean vowel duration as a function of vowel length and onset size (subsets A and B, left) or coda size (subsets C and D, right).

A final comment about the duration of empty onsets and the effect of empty codas on vowel length in this experiment might be in order. In these respects, we find the results to be very similar to those found in the pilot experiment. Long vowels before empty codas (figures 8 and 4, to the right of the line) are about 25 ms longer than long vowels before monosegmental codas in both experiments, confirming Quené’s (1989) claim that word-final long vowels are longer than such vowels in closed syllables (see section 2.3.2). Empty onsets (figures 7 and 3, “0”) have a duration of roughly 80 ms in both cases (averaged over long and short vowels in the second experiment). In the control experiment, as well as in the pilot experiment, the duration of the empty onset is not comparable to the duration of either “filled” onsets or “filled” codas. Again, a tendency to group onsets with the empty onset is not revealed by these production data.

In conclusion we can say that this experiment has revealed two important facts about duration increase in onset and coda clusters of different sizes. Firstly, the segmental content of the clusters and the positioning of these segments in the cluster does not seem to affect the general tendency to increase in duration with the addition of segments that we find in codas, but, more importantly, also in onsets. The results of the two experiments described in this chapter are very similar in this respect, though the identity of the coda segments and the position of the segments in the second experiment differed with respect to the first experiment. In the second experiment we used coda clusters that were the mirror images of the onset clusters. Yet we did not find any difference between those clusters as far as the duration increase under the influence of the addition of extra segments is concerned. Hence, this experiment shows that onset duration behaves exactly like coda duration in speech production. Onset duration is not invariant with respect to coda duration, which means that we must reject the Weak Hypothesis. Secondly, the usage of exact mirror clusters in the second experiment seems to have sharpened the compensatory effects that we expected to find between nuclei and codas. Such compensatory effects between onsets and nuclei are negligible (see figure 9).

2.5 General discussion

The experimental results presented in this chapter clearly show that we cannot find an explanation for the weightlessness of the syllable onset in speech production.[9] The durational invariance that we hoped to find in the

---

[9] Browman & Goldstein (1988), though, have traced an articulatory property of onsets that may lead to a possible explanation. They introduce the C-centre of consonant groups which is defined as the mean of the midpoints of all articulatory gestures that are used in the consonant cluster. In their experiments they find the duration from the onset’s C-centre to the VC transition to be constant, whatever the nature and size of the onset. For coda clusters this does not hold, coda C-centres expand with respect to their own starting point, which is also the VC transition. Browman & Goldstein claim that this is an organizational

onset is not reflected in the data. Figures 3 and 7 clearly show that onsets of different sizes have different durations. These results show that the tendency of invariant cluster durations reported by Nooteboom & Cohen (1988) must have been based on idiosyncratic properties of the spectrograms they present. They introduce the possibility that speakers tend to keep the duration of consonant clusters constant irrespective of the number of segments in that cluster. This possibility cannot be confirmed by the experimental data we have presented here.

In section 2.1 above we claimed that we should not expect to find a rigid form of phonetic invariance of onset duration and that relative durational invariance of the onset with respect to the coda would constitute enough evidence for the phonological weightlessness of the onset. Unfortunately, even this relative form of invariance cannot be found in our data. The results found for coda duration are comparable to the results for onset duration in all cases, as can be seen when we compare figures 3 and 4, and figures 7 and 8. Reprocessing the data using a slightly different definition for the right hand boundary of the weightless part of the syllable does not help. Hypothesising that the first part of the syllable up to the sonority peak (which rises in sonority) is weightless, we tested whether this syllabic division yielded relative durational invariance for the relevant part. Figures 5 and 6 show that it did not.

If anything, our results rather constitute evidence in favour of weightless rhymes than weightless onsets. We have found that both onsets and codas increase in duration when segments are added. Hence, they are not durationally invariant and, by our definition, not weightless. When we look at the rhyme, however, we see that there are significant compensatory effects that reduce the impact of duration increases in the

---

difference. Onsets are linked to the vowel while codas are sequentially ordered and may have their own time frame. They relate this observation to the fact that onsets are always weightless while codas are not, but limit themselves to saying that the two phenomena are correlated. In our view, an explanation for the onset's weightlessness through C-centres can indeed only be found if C-centres are perceptually relevant. If some onset property is to play a role in the phonology, it must be audible (like in our initial hypothesis where onsets were deemed to have invariant durations, a fact that must be audible). As yet, durations of onset clusters with invariant C-centres may vary greatly. Some other perceptible phenomenon related to C-centres must then be at play to link the invariance of C-centre location to onset weightlessness. Browman & Goldstein define the C-centre in an articulatory way (production), but propose that it is related to the perceptually defined P-centre (cf. section 4.3 and references cited there). To check whether P-centre and onset weightlessness are related we first need to gather more knowledge about the perception of syllabic duration (see chapter 3). Then we can verify whether the correlation between fixed C-centre location and phonological weightlessness is explanatory or coincidental.

coda: the total increase in rhyme duration is dampened by the shortening of the vowel that occurs when a segment is added to the coda. For the onset this effect is less significant. Therefore, it is the rhyme rather than the onset that is relatively invariant in duration. This is contrary to what we would expect from a phonological point of view. It is clear that we were on the wrong track in our assumption that the phonetic correlate of phonological weight could be found in the actual physical duration of the segments in speech production.

The results found in this chapter are confirmed by Waals (1996). In an extensive production experiment she determines the amount of segmental compression (ie. compensation in duration) in an impressive number of Dutch word onsets. Her final conclusion is that single onsets have a mean duration of 150 ms, binary clusters are 200 ms on average, while clusters with three segments are roughly 250 ms long.[10] For coda clusters Waals (forthcoming) reports 80, 140 and 170 ms as durations for one, two, and three segments respectively.

Further confirmation for these results comes from our own processing of data that were independently gathered by Hofhuis (1993).[11] From this elaborate set of data we took (nonsense) words that more or less resembled the type of words presented in (1) and (2). The words we used are given in (3).

(3)	laat	'late'	taal	'language'
	slaat	'(he) hits'	taals	nonsense word
	splaat	nonsense word	taalts	nonsense word

The mean cluster durations for the relevant onsets and codas were calculated from the data file. The results are similar to those found by Goedemans & van Heuven (1993) and Waals (1996). Onset duration grows significantly with cluster size: (one-way anova) $F(2,33)=162.3$ ($p<.001$), and so does coda duration: $F(2,33)=70.7$ ($p<.001$). A table with the mean durations taken from these data can be found in appendix A (Table IV).

The general conclusion must therefore be that there is no temporal asymmetry present in the speech wave from which the listener can derive a difference in phonological weight. To find a phonetic explanation for

---

[10] Waals' exact values for the clusters that we have considered above are: 159 for /s/, 192 ms for /st/ and 253 for /str/.

[11] Many thanks to Elise Hofhuis for allowing us to work with her data. We undertook this exercise to enlarge the data set, thus increasing the support for our claim that Dutch syllable onsets are not durationally invariant.

the weightlessness of the syllable onset we must obviously look beyond speech production. As Gimson (1975) notes:

“Clearly, whenever it is possible to establish the boundaries of sounds or syllables, it will be possible to measure their duration by means of such traces as are provided by oscillograms or spectrograms. Such delimitation of units, in both the articulatory and the acoustic sense, may be difficult... But, even when it can be done, variations of duration in acoustic terms may not correspond to our linguistic judgments of length... This distinction between measurable duration and linguistic length provides another example of the way in which our linguistic sense interprets from the acoustic material only that which is significant.”(Gimson 1975:24)

It is indeed often noted that human listeners do not perceive the speech signal as a machine would. The way in which our hearing mechanism is constructed may influence our perception of speech. This might result in our hearing differences that are not physically present in the speech wave. A simple example is the following. The sensitivity of the ear to intensity depends on the frequency of the signal (cf. Handel 1989). A sound of 40 dB is clearly audible at 1000 Hz, but at 100 Hz we can hardly detect it; we hear a difference in intensity that is not there. It might well be the case that, for some reason or other, there is a difference in the perception of duration between the onset and the rhyme of a syllable. This difference in duration perception could be the explanation for the weightlessness of the onset. In the next chapter we discuss a number of perception experiments that were conducted to test the hypothesis that such a difference in duration perception exists between onset, nucleus and coda, and that the phonetic explanation for the absence of phonological weight in onsets is contained in this difference.

4 Exploratory Psychophysics[1]

4.1 Introduction

The experimental results that were presented in chapter 3 show that weight and duration perception are related. We have demonstrated that the perceived duration of the syllable onset, which is phonologically weightless and, hence, can bear no mora, is relatively invariant with respect to the perceived durations of nuclei and codas, which both may bear moras. This does not mean, however, that we can conclude our research. On the contrary, we must now take our research question one step further and ask ourselves why duration perception in onsets is so poor. Relatively inaccurate duration perception cannot be the ultimate cause of weightlessness: there must be some more basic effect that causes this inaccurate duration perception. We conjecture that the difference in the duration perception of onsets, nuclei and codas might be caused by a phenomenon of a more general, psychophysical nature. In this chapter, therefore, we address the question what that phenomenon could be. The answer to that question should give us the real reason why onsets are weightless.

Two possible psychophysical causes for the difference in duration perception of onsets, nuclei and codas come to mind. Firstly, there might be a physical property of onsets in general that jams the (neuro-) physiological "mechanism" with which we measure duration. The related hypothesis would be: 1) *variation in onset duration is poorly perceived because onsets have certain characteristics that make them impervious to correct duration perception*. A characteristic we could think of is the rising intensity that we find in all onsets. Secondly, we might assume that subjects only start paying attention to duration after a certain salient point in the syllable. If this point occurs to the right of the onset consonants it is only logical that variation in their duration is perceived rather poorly. This is the first possibility that we will explore in the experiments below. The hypothesis that we will test is the following: 2) *perceived onset*

[1] The results of the experiment described in section 4.2 have been published as the second half of Goedemans & van Heuven (1995).

duration is relatively invariant since the listener's attention to duration is triggered by a certain salient point in the syllable which comes after the onset consonants.

One likely candidate that could serve as a salient point in the syllable is the CV transition (or PIVOT), which according to Dogil & Braun (1988) serves as an anchor point in the perception of syllables. Ohala & Kawasaki (1986) point to the magnitude and the rate with which several stimulus parameters vary near the CV transition as the main cause of its saliency. Indeed, CV transition characteristics like a rapid rise in intensity and swift formant transitions belong to the most salient events occurring in syllables. Ohala & Kawasaki conclude that these salient prevocalic events constitute an ideal timing mark for the synchronisation of the segmental and the prosodic stream. We propose that it is not far fetched to assume that listeners start paying attention to duration after this timing mark, ignoring duration variation in the consonantal cluster that comes before it.

This timing mark occurs later in the syllable as the onset lengthens. As a consequence, we cannot provide independent evidence for the CV transition as the duration perception trigger. Inducing variation in the location of the CV transition without changes in the duration of nucleus or onset is nearly impossible (but see the discussion in 4.4.3). We already know that lengthening of the onset does not drastically alter the perceived duration, but we cannot find out whether that is the result of the CV transition triggering the start of accurate duration perception only after the (lengthened) onset, or whether it is due to some other property of the onset or the point at which the onset ends and the vowel begins (i.e. another trigger). The only possible strategy here is to check all other possible explanations, and when they fail, adopt the delayed duration perception hypothesis (with the PIVOT as the trigger) as the most likely alternative. More will be said about this PIVOT option below.

Besides the PIVOT, two other candidates for the trigger come to mind. Firstly, the intensity maximum is a salient point that usually occurs somewhat later than the PIVOT. Fortunately, this is a trigger of which the location in the syllable can be altered without altering the segment durations. If the intensity maximum triggers the duration perception "mechanism", we expect perceived durations to shorten when this maximum occurs later in the syllable. The internal chronometer starts later, and only the duration of the syllabic part after the trigger is perceived accurately, or in other words, when the trigger moves to a point later in the syllable, the lengthening of the part to the left of it will not be readily perceived while the shortening of the remaining part will, leading to a decrease in perceived duration.

A related third candidate is the *P-centre*, which can be defined as the point at which a stimulus rhythmically occurs (Pompino-Marschall 1989, 1991; Howell 1988; and Scott 1993, among others). In spite of all these studies, the P-centre remains an elusive phenomenon. To get an idea of what it represents it helps to first envisage a series of syllables of which all starting points are equidistant in time (say at 500 ms intervals). It has been shown in the past that listeners do not perceive these onset-isochronous sequences of syllables as being regularly spaced in time (Fowler 1979). If we ask subjects to “correct” these timing irregularities by aligning the syllables to the beats of a metronome, we find that they invariably place them such that the metronome beats occur at points well after the beginning of the syllable.[2] For each syllable, the metronome beat marks the psychological moment of occurrence. This point, sometimes also referred to as the perceptual beginning of the syllable, defines the P-centre (Morton, Marcus & Frankish 1976). Hence, for this sequence of syllables which perceptually occur at regular time intervals, the P-centres are equidistant. In figure 1 the candidates are schematically represented.

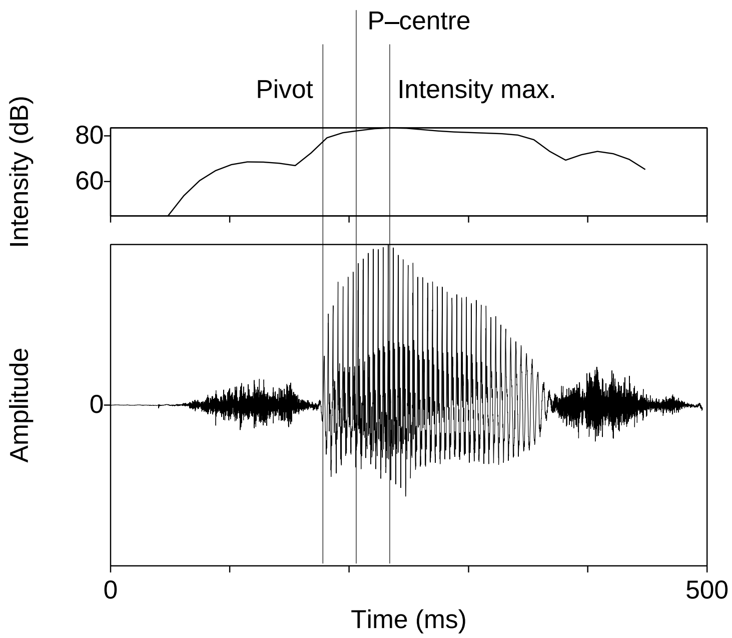


Figure 1: schematic representation of the PIVOT, P-centre and intensity maximum (in *saas*).

[2] The P-centre usually occurs near the CV-transition. For more details on the location of the P-centre, and the metronome beat experiments that can be used to find them, see the references given at the beginning of this paragraph.

Apparently the P-centre is an important concept in the timing of syllables and must not be overlooked when we consider possible explanations for the weightlessness of the syllable onset. It sounds logical that listeners only start hearing duration differences *after* this psychological moment of occurrence of the syllable. Hence, if our assumption that a trigger is needed for faithful duration perception is correct, we expect perceived durations to shorten as the P-centre occurs later in the syllable. Since the location of the P-centre is dependent on the shape of the intensity curve (Pompino-Marschall 1989) we cannot suffice by only checking whether the intensity peak is the trigger we are looking for. We must also check the P-centre. Afterwards we can determine whether the position of the P-centre or the position of the intensity maximum correlates best with perceived duration, and determine which of them is the most likely trigger. More on the nature of the P-centre, its role in the psychophysical experiments in this chapter and its relation to the intensity maximum will be said below, after the presentation of an experiment in which we explore the possibility of the intensity maximum as the duration clock trigger. Discussion of this experiment follows in the next section.

4.2 The relevance of intensity peak position for duration perception

In this section we discuss two experiments that were conducted to verify whether the position of the intensity peak in the syllable affects the perceived duration of that syllable. We employ the two experimental strategies that were introduced in the previous chapter. First we will present an experiment in which we compare a range of stimuli with variable intensity peak positions to a fixed reference stimulus, using the 2I-2AFC paradigm. After that we will present a control experiment in which we use the same stimuli in a duration adjustment task. The stimulus syllables will be identical in every respect apart from the position of the intensity peak, which will be shifted through the syllable with regular intervals. All speech stimuli will be matched by noise stimuli of which the intensity envelopes will be manipulated in a similar fashion. Through usage of these noise stimuli we may be able to determine whether the suggested influence of the intensity peak on duration perception is purely a speech phenomenon or whether it is a more general property of sound perception.

4.2.1 Pairwise comparison of stimuli with different intensity peak positions

In this experiment we test whether the position of the intensity peak influences the perceived duration of a given stimulus. Our expectation was that perceived duration should shorten as the intensity peak moves to the right. To verify this we take a reference stimulus (Sr) with a neutral value for peak position, and test in a 2I-2AFC task whether comparison stimuli (Sc) with peaks occurring towards the right edge of the syllable are judged shorter than the neutral reference stimulus, and whether stimuli that have peaks towards the left edge are judged longer than this neutral stimulus. There are at least two possible reference stimuli we can use for this task. The first is a stimulus that has a *flat* intensity envelope without a peak (Sr-f[lat]), and the second has an intensity *peak* which is located exactly in the middle of the syllable (Sr-p[eaked]). We do not know beforehand which of these two reference stimuli will yield the best results, so we use both. For both reference stimuli we expect the number of "comparison stimulus is longer" judgements to decrease as the peak in the comparison stimuli moves from the middle of the syllable to the right (assuming that the listener's attention is drawn to duration perception after the intensity peak has occurred) while the number of "Sc longer" judgements should increase as the peak moves from the middle to the left.

4.2.1.1 Stimuli

In most of the experiments that were described in chapter 3, the Dutch word *mam* ('mother') was used to generate the speech stimuli. To make sure we do not lose the connection with these former experiments, we will continue experimenting with this word. We used the original stimulus that was also used in the previous experiments. In this stimulus (synthesized from diphones, LPC 10, 5 formants/Bandwiths, 10 ms frames, speaker HZ) the durations of the three segments had been adjusted so that they were all 100 ms, and the linear rise and fall time had been set to 30 ms to avoid disturbing clicks. The second original stimulus was a 300 ms burst of white noise of which rise and fall time were adjusted in the same manner. The intensities of these stimuli were level, and the intensity of the noise burst was modified until it sounded as loud as the speech stimulus. To generate the set of comparison stimuli we further adjusted the intensity envelopes of the speech and the noise stimulus. Stimuli with intensity peaks were generated according to the scheme in table I. The two original files (speech and noise), and two that had intensity peaks at 150 ms (the exact middle of the syllable) were used as neutral reference stimuli (Sr-f and Sr-p, respectively). The resulting set of 28 stimuli contained the 14 stimulus types in table I for both speech and noise. In appendix C, figure 1, the schematised intensity envelopes of the boldfaced stimuli in table I are presented.

Table I*: stimuli for the pairwise-comparison experiment with their intensity peak positions (with respect to syllable beginning). Stimuli in bold are illustrated in appendix C, figure 1.*

	Reference Stimuli		Comparison Stimuli (Sc)
	Sr-f	Sr-p	
Peak Position in (ms)	**none**	**150**	40 60 80 **100** 120 140 160 180 **200** 220 240

On auditory inspection of the stimuli it appeared that both the Sr-f for speech and noise sounded longer than their Sr-p equivalents (they were at least louder, which may be caused by the greater amount of energy that is found under the intensity curve in these stimuli, see appendix C, figure 1). A small pilot experiment was conducted to determine the absolute perceived duration for the two different types of Sr. We used the adjustment method described in section 3.1 to determine the perceptual duration difference between these stimuli. Five subjects adjusted a white noise burst until it was perceptually equal to one of the reference stimuli. The stimuli were repeated 6 times. The results showed a mean adjusted duration for Sr-f that was 30 ms longer (244 ms) than the adjusted duration for Sr-p (213 ms). On the basis of these results Sr-f was shortened to 270 ms by deletion of a 10 ms frame in each segment (and 3 arbitrary frames from the noise stimulus). The parameter files were converted to sampled data files using LPC synthesis for VAX/VMS (cf. Vogten 1984).

4.2.1.2 Subjects and procedure

Forty-six paid Dutch subjects participated in the experiment. Their ages varied between 18 and 40, and none of them reported any hearing difficulties. The experiment was conducted in a sound attenuating booth. Stimuli were presented in pairs to the subjects according to the scheme in (1). Each combination represents a block of 13 pairs in which the 12 Sc's from table I are compared with one of the Sr's, while Sr-p was used as an extra comparison stimulus in the 13th pair of each block.

(1)	Sr-f (speech)	- Sc (speech)	Sr-p (speech)	- Sc (speech)
	Sr-f (noise)	- Sc (noise)	Sr-p (noise)	- Sc (noise)
	Sc (speech)	- Sr-f (speech)	Sc (speech)	- Sr-p (speech)
	Sc (noise)	- Sr-f (noise)	Sc (noise)	- Sr-p (noise)

The blocks were presented in this order, but the internal order of the stimuli was random. The variation in the position of Sr and Sc should balance the Time-Order Error (see van Heuven & van de Broecke 1982 and chapter 3 for reference to this phenomenon). In total 107 pairs (8 x 13 + 3 practice stimuli), were presented consecutively to the subjects via an Iris Indigo workstation (experimental set-up identical to that described in section 3.4.1). A pause of 500 ms came between the members of a pair. The subjects had to decide which member was the longer of the two by pressing a green (first longer) or a red (second longer) key on the keyboard. They also had the possibility to repeat the trial by pressing the spacebar. Once a response was given, the computer automatically turned to the next trial. The responses were recorded and stored on disk.

4.2.1.3 Results and discussion

We calculated the percentage of subjects that judged Sc to be longer than the neutral reference stimulus (Sr) for each value of peak manipulation. If the trigger hypothesis is correct, and the intensity peak is the trigger, early peaks should trigger many “longer” judgements, while late peaks should not. Therefore, if we plot these calculated percentages as a function of peak position, we expect the psychometric curves to run from approximately 100% for the leftmost peak to 0% for the rightmost one. For the pairs that consisted of two equal reference stimuli subjects should not be able to pinpoint the longer one, and score at chance level (50%). These expectations are not borne out by the facts, as we can see in figure 2 in which the results for the first part of the experiment (with Sr-f) are given (collapsed over the two time orders).

In this figure we see the percentage of subjects that judged the stimuli with the shifted intensity peaks to be longer than the reference stimulus with the flat intensity envelope for each peak position in the noise and speech stimuli. As we can see, there seems to be no *clear* effect of intensity peak position on duration perception. We also observe that speech and noise behave similarly in this respect. The lines through the data points are more or less flat: no matter what we do with the position of the intensity peak, the perceived duration does not drastically change.

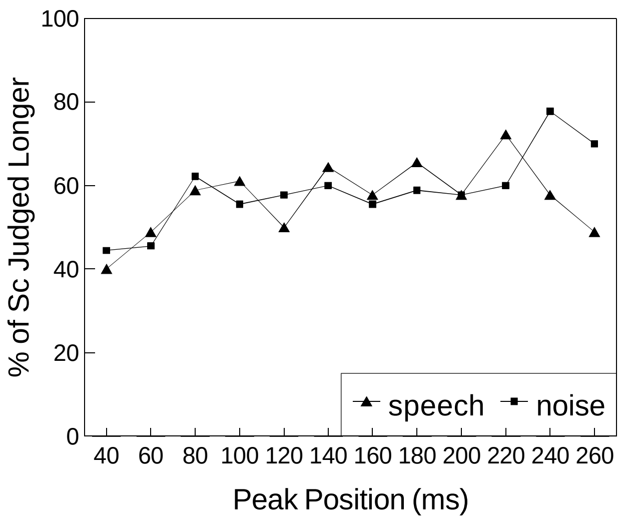


Figure 2: percentage of “longer” judgements for Sc as a function of peak position (for Sr-f).

Judgements vary just above the 50% chance line for most values of the peak position variable.[3] This means that subjects have a bias for Sc when they choose the longer stimulus from an Sr-Sc or Sc-Sr pair, which is unexpected. If the intensity peak position has no effect on duration perception, all stimuli should have equal perceived durations, and subjects should score at chance level in all cases. An explanation for this deviation from chance levels is readily available though. Note that we shortened Sr-f because it sounded longer than Sr-p. Probably the 30 ms with which we shortened Sr-f on the basis of the small pilot experiment was too much, the result being an Sr-f which is now perceptually a little shorter than Sr-p.[4] As we have noted, figure 2 shows there is almost no perceived duration difference between most of the comparison stimuli.

---

[3] The original data showed one aberrant data point for the 260 ms speech peak at about 20%. It appears that this is caused by the position of one of the repetitions for this stimulus in the ordering of the stimulus pairs. It was located exactly after the transition from Sr-Sc ordered noise to Sc-Sr ordered speech. The same was true for the Sr-p stimulus list and these data showed exactly the same aberrant point. It must be the case that the position this stimulus had in the list caused it to be underestimated in duration, but we do not know why. Because it was so clearly an experimental artefact, we felt justified in deleting this stimulus from the data set and calculating the percentage for speech 260 on the basis of one stimulus only. We repeated this procedure for Sr-p below.

[4] This error in the amount of shortening needed for Sr-f may well have been caused by the relatively small number of subjects that were used in the pilot.

Hence, the bias affects all of them. Remember also that we added Sr-p for reference. Its percentage of “Sc longer” judgements is 60 for both speech and noise. We observe no obvious differences between the comparison stimuli and Sr-p. Hence, the perceived durations of all Sc’s are probably equal to that of Sr-p (as we will see below in the second part of this experiment). This means that Sr-p and all Sc’s are perceptually longer than Sr-f, and it is only logical that subjects show a bias towards longer Sc’s when they are compared with Sr-f. This bias is of little consequence to the expected trend, however. If an effect of intensity peak on duration perception had been present in our data, the result of the shorter Sr-f would have had no great effect on the shape of the tilted line from 100 to 0% that we expected to find. Some higher percentage scores in the middle region of the line (curving it up a little) would have been the only result of this shorter Sr-f. As it is, we do not find such a tilted line at all. If anything, the small effect that seems to be there is the reverse of what we expected. We detect a gently rising slope, which indicates durations are judged marginally longer as the peak moves to the right: correlation coefficients are 0.43, $p<.05$, for speech and 0.67, $p<.001$, for noise. An explanation for this opposite trend is not available.

The time-order error, which we tried to control by using two stimulus orders, is absent in these results. Underestimation of the final stimulus should lead to less than 50% “Sc is longer” judgements in the Sr-Sc order while the “longer” judgements for Sc should exceed 50% for the Sc-Sr ordering. What we find is 58.8% for Sr-Sc and 58.6% for Sc-Sr. Hence, there is no TOE effect for the Sr-f data. Perhaps the difference in perceived duration between Sr-f and Sc has overshadowed this TOE effect. Since we tried to keep the TOE effect out of our data in the first place (albeit, in a different manner), we are not bothered by its absence.

The explanation for why “longer” judgements for Sc exceeded the 50% chance level in figure 2 crucially hinged on the assumption that Sr-p is equal in perceived duration to the Sc’s (hence marginally longer than Sr-f). We have assumed that this is the case since the only difference between the comparison stimuli and Sr-p is the position of the intensity peak, which does not seem to influence perceived duration. We have not presented any evidence, however. In figure 3 we find such evidence. It shows the results for the Sr-p part of the experiment in the same fashion as we to presented the Sr-f data.

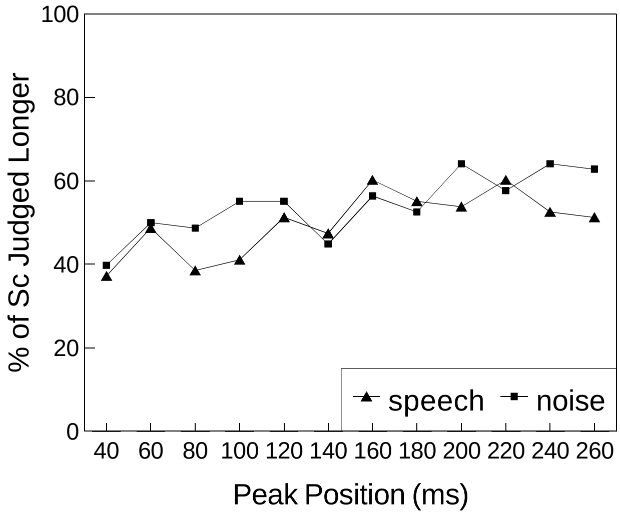


Figure 3: percentage of “longer” judgements for Sc as a function of peak position (for Sr-p).

The data points in this figure represent the percentages of subjects that judged the stimuli with the shifted intensity peaks to be longer than Sr-p for each peak position in the noise and speech stimuli. Contrary to what we find in figure 2, judgements centre around the 50% line, indicating that there is no difference in the perceived durations of the shifted-peak stimuli and Sr-p. This means our reasoning with respect to the bias for the stimuli when compared with Sr-f was justified.

As far as the TOE is concerned, the average judgement percentage is close to 50. Our suggested explanation for the absence of TOE in the Sr-f data was that the difference in perceived duration between the reference and the comparison stimuli in that case overshadowed the TOE effect. Since we find an average of 50% for the Sr-p data, there is probably no such overshadowing factor present in this case. Hence, we should be able to trace the TOE effect. Indeed we find a small effect: percentages for “Sc longer” judgements are 49.6% for Sr-Sc and 53.1% for Sc-Sr, a slight tendency towards underestimation of the final stimulus (or overestimation of the prefinal one), exactly as is reported in the literature (note that we cannot calculate the TOE in ms here, like we did in chapter 3, because we cannot determine the PSE in this case). The fact that we do find a TOE effect for the Sr-p data strengthens our conviction that its absence above was caused by the shorter perceived duration of Sr-f.

In every other respect figure 3 is comparable to figure 2. Again we observe that the position of the intensity peak seems to have a small but reversed effect on the perceived duration. Percentage scores are comparable for every peak position in both the noise and the speech stimuli. The only effect that can be found in this figure is again a gently rising slope in the lines connecting the data points. Hence, also for Sr-p there is a slight tendency towards longer judgements as the peak moves to the right, contrary to our expectations. Correlation coefficients are 0.52, $p<.01$, for speech and 0.63, $p<.001$, for noise. Ignoring this small unexplained effect we may conclude that, to all intents and purposes, all comparison stimuli used in this experiment have more or less equal perceived durations. We have not found the predicted effect of intensity peak position on perceived duration. Therefore, we might feel forced to abandon the hypothesis that the intensity peak triggers a more accurate mode of duration perception, even if we do not ignore the small effect referred to above (which runs counter to the predictions of our hypothesis). However, we cannot do so before we have presented the second experiment that we conducted simultaneously, using the same stimuli. If the desired effect is also absent in the results we obtain in this experiment, we may safely reject the intensity peak as a candidate in the trigger hypothesis.

4.2.2 The relevance of intensity envelope in a duration adjustment task

In this experiment we checked the behaviour of the stimuli that were used in the previous experiment in an adjustment task. In such a task, subjects adjust the duration of a comparison stimulus until it is perceptually equal (in duration) to a reference stimulus (see chapter 3 for a description of this paradigm). Abiding with the intensity-trigger hypothesis, we might expect subjects to adjust the comparison signal to a shorter duration as the intensity peak moves to the right. However, if the results of the previous experiment are to be taken seriously, we should expect adjusted durations to show little variation, even tending towards a rising trend.

4.2.2.1 Method

In this experiment the comparison stimuli described in section 4.2.1.1 were used as *reference* stimuli, while an adjustable burst of white noise served as the comparison stimulus in each case. The noise burst had exactly the same characteristics as the Sr-f stimulus in the previous experiment, only its duration was variable and its offset was, for technical reasons, abrupt.[5] We chose to use the noise burst as the comparison signal for both the speech and the noise stimuli, instead of re-using for speech the periodic signal that we used for the *mam* stimuli in chapter 3. There, the reason for using the periodic non-speech signal was that it should encourage subjects to include the nasal consonants when estimating syllabic duration. This time, though, we want to compare the adjustments for speech stimuli to the adjustments for noise stimuli. The noise stimuli obviously need a noise comparison signal, and we feel that the signals to be adjusted should be the same to be able to make a fair comparison of speech and non-speech. The neutrality of noise with respect to the fully periodic *mam* should help subjects to make reasonably correct adjustments, so we do not expect the results to suffer greatly from this change of comparison signal.

Subjects were 40 native speakers of Dutch, with ages between 18 and 40, and without any self-reported hearing deficiencies. They received payment for their participation. The experiment was held in the same type of sound-proofed booth as the previous experiments. The experimental set-up was identical to that described in section 3.2.1. Each reference stimulus occurred twice in the stimulus list, and the pairs were presented in random order. Within each pair, only the order Sr - Sc was exploited to reduce the workload on the subjects. Subjects adjusted the duration of a the white noise burst, which was presented 300 ms after the reference stimulus (offset to onset). Reference-comparison pairs were repeated with 1000 ms silent intervals. The duration of the noise burst could be adjusted during this interval. The effect of the subjects' adjustments was audible in the comparison signal during the next repetition of the pair. Subjects were allowed to make as many adjustments as they liked. Once they were satisfied that the comparison and the reference signal were of equal duration, they could store the result on disk, after which the next trial was initiated. At the onset of each new trial the duration of the comparison stimulus was 0 ms.

4.2.2.2 Results and discussion

In view of the rather poor performances of some subjects in the adjustment experiments that were presented in chapter 3, we decided to test the subjects for internal consistency in their adjustments, like we did in section 3.3.2. Eight subjects correlated below r=0.3 in their first and second adjustment of the same stimulus and were taken from the data set.

[5] This comparison stimulus also closely resembles the noise burst that was used in the adjustment experiments described in chapter 3.

As noted above, we must not expect perceived duration to decrease as the intensity peak moves to the right. In the light of the results found in the previous experiment we do expect the results to show no effect of intensity peak position on adjusted duration, or even slightly longer adjusted durations as the peak moves to the right. In figure 4 we see that these expectations are justified. In this experiment, however, the reverse trend that we observed in the previous experiment seems to be absent.

We find here the mean duration that the noise burst had when subjects judged their adjustment of this comparison signal to be such that its duration was equal to that of the reference stimulus. Mean adjusted durations (y-axis) are given for each value of peak position (x-axis) and broken down by stimulus type (speech stimuli: bottom line, noise stimuli: top line). For the noise stimuli we observe that the line through the data points is nearly straight and horizontal. This time, there is clearly no effect of intensity peak position on perceived duration. Had this effect been there, the line should have been tilted, the adjusted duration for the 260 ms peak being significantly lower than that for the 40 ms peak. A one-way anova shows that this significant effect is absent: adjusted duration by peak position for noise $F(11,752)=0.9$, ins. The line for the speech stimuli is also horizontal and nearly straight, apart from two spurious peaks. Again a significant effect of peak position is absent: $F(11,752)=1.2$, ins (see also footnote 7 below).

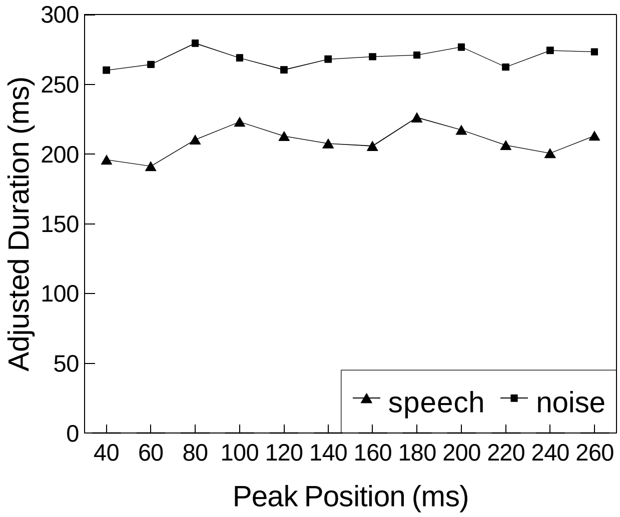


Figure 4: mean adjusted durations as a function of peak position for speech and noise stimuli.

The two peaks that stand out in the speech line, are located at 100 and

180 ms. We do not view the fact that these peaks occur near the CV and the VC transition points as a coincidence, but we have no explanation to offer for this phenomenon. Further note that the rising trend that we observed in the previous experiment is absent here. We conclude from this experiment that speech and non-speech again behave similarly with respect to the influence that the position of the intensity peak has on perceived duration. The bare fact is that no difference in perceived duration can be induced by movement of the intensity peak, neither in syllables, nor in noise.

One striking difference between the speech and non-speech data points in figure 4 has not yet been addressed. Adjusted durations for speech are just above 200 ms most of the time and never exceed 240 ms, while the data points for non-speech centre around 275 ms. Compare the adjusted durations for speech to the adjusted durations for speech shown in chapter 3, figure 2. The adjusted durations for 300 ms *mam* syllables are almost identical to the adjusted durations we find here. We conclude from this that it did not matter after all whether we used a periodic or a noise signal for the comparison stimulus. The adjustments for the 300-ms reference syllable are about 200 ms, irrespective of the type of comparison signal. We also consider this resemblance to be evidence for the validity of the results found in the second experiment in chapter 3.

An even more important conclusion can be drawn from the difference between the adjustments for speech and non-speech we find here. If the noise burst is adjusted to 200 ms to achieve perceptual equality with a 300 ms syllable, and *the same noise burst* is adjusted to 275 ms on average to achieve perceptual equality with 300 ms white noise stimuli, this must mean there is a difference in the perceived durations of the 300 ms syllable and the 300 ms noise burst. The noise burst sounds about 75 ms longer than the syllable. This means that we have found an answer to the puzzling underadjustments that we found in the first two experiments in chapter 3. There the adjustments for the *sas* and *mam* syllables were systematically too low. We know now that this phenomenon is caused by the fact that 300 ms noise comparison signals are perceptually longer than 300 ms syllables. A signal that inherently sounds longer must be underadjusted to have a duration that is perceptually equal to that of other signals like speech (compare chapter 3 footnote 3). The fact that the mean adjustment for noise is 275 and not 300 ms is possibly due to the difference in intensity envelope that also caused Sr-f to sound longer than Sr-p in the previous experiment. Remember that the reference stimuli in this experiment all have intensity peak envelopes (like Sr-p), while the comparison stimulus has a flat envelope (like Sr-f). The 25 ms difference between adjustment and actual duration resembles the 30 ms correction we applied to Sr-f to make it sound as long as Sr-p (which appeared to be a little too much in hindsight).

The general conclusion we may draw from the results of the two experiments presented in this section is that the position of the intensity peak does not influence perceived duration. Speech and non-speech are not different in this respect. In both speech and noise the perceived duration is identical for all the stimuli we incorporated in the experiments, though the location of the intensity peak was varied over almost the entire timespan of the stimuli. We may, therefore, safely rule out the intensity peak as the trigger of more accurate duration perception. If our initial hypothesis is correct, and if indeed there is a salient point in the syllable at which duration perception improves, then this point is definitely not the intensity peak. Remember, though, that in section 4.1 we introduced a candidate that was related to the intensity peak, namely the P-centre. However, it is most likely that the experimental results presented above also disqualify the P-centre as a trigger candidate. In the next section we give some more background on the nature of P-centres, and we explain why this P-centre can no longer be considered a likely trigger on the basis of the absence of an effect for the intensity peak. We then present an experiment in which we try to find evidence for this position.

4.3 The role of the P-centre

As noted in section 4.1 above, the P-centre is the perceptual moment of occurrence of an auditory stimulus. It is defined by the position of that stimulus in a rhythmic sequence. If an intrusive syllable in a series of beats sounds rhythmical, the P-centre of that syllable is located exactly where the beat occurred that is replaced by the syllable. Alternatively, we can find the P-centre by having subjects tap out (with a pen on a desk) the rhythm of a series of identical syllables that are equally spaced in time. The P-centres should then be located at exactly the same point in each copy of the syllable. Hence, since the syllables occur at regular intervals, so should the P-centres. The taps will reveal the P-centre of the syllable in question, which is usually located somewhere near the CV transition. In this section we focus on the occurrence of P-centres in syllables, but other acoustic signals, like our noise stimuli, also have a P-centre, as we will see below.

Attempts to define the P-centre through the phonetic properties of the syllable in which it occurs have revealed some properties that influence P-centre location, but the results do not yield a single procedure by which we can calculate the position of the P-centre for any given syllable. Recourse to experimentation must still be taken to determine its exact location. Marcus (1981), for instance, claims that the location of the P-centre depends on the durations of the onset and the rhyme according to the following formula:

(2) $P_c = 0.65 \bullet C_1 + 0.25 \bullet VC_2 + c$

in which *c* is a constant that seems to depend on external factors. The value of *c* can only be found through experimentation. Further investigations (Pompino-Marschall 1989, 1991; Howell 1988; and Scott 1993) have shown that other factors, like energy distribution, the shape of the intensity curve, and maybe even F0-movement, also play a role in the determination of the P-centre location. Yet, to date the complete set of determining factors has not been found. The best way to find the P-centre is still to conduct a beat alignment or a tapping experiment.

The importance of the P-centre research mentioned above is that it shows that the P-centre cannot be manipulated without altering some other syllabic property, like duration or intensity envelope. This seemed to complicate our task at first. There was no way in which we could independently move the P-centre through the syllable and test whether these movements had any effect on the perceived duration of that syllable. In the light of the previous results, however, a new strategy emerges. Since the P-centre depends heavily on the shape of the intensity envelope, we expect the P-centre to vary noticeably in the stimuli we created for the two previous experiments. On the basis of what Pompino-Marschall (1989) and Scott (1993) report, we may expect the P-centre in our stimuli to move through the syllable, in proportion to the intensity peak manipulations. However, we already know that the perceived durations of these manipulated syllables are equal. Hence, our expectation with respect to the candidature of the P-centre for the trigger position has to be changed. If the P-centre moves, and yet the perceived duration of the stimuli does not change, then the previous experiments have shown that neither the intensity peak nor the P-centre is the trigger we are looking for. What we must do now, of course, is experimentally determine whether the P-centres in our stimuli do indeed vary with the position of the intensity peaks. The experiment we conducted to find the P-centre locations is described below.

4.3.1 Determining P-centre locations

In this experiment we determine the P-centres of the stimuli we used in the previous experiments. The location of these P-centres is partly determined by the position of the intensity peak, and hence, we expect them to move trough the syllable in proportion with peak movement. If this experiment does indeed show P-centre variability, we can discard the P-centre as a candidate for the onset point of proper duration detection. We can do this because we already know that our stimuli yield no perceived duration differences. If the P-centre is the durational trigger, perceived durations should shorten as the P-centre occurs later in the syllable (cf. also the reasoning for intensity peak). Hence, the absence of a duration difference in our stimuli could also be caused by an invariant P-centre location. Should the present experiment reveal that the P-centre remains constant, in spite of the intensity peak shifts, we must design a new experiment in which we move P-centres to test whether such shifts yield perceived duration differences.

4.3.1.1 Method

As has been mentioned above, the stimuli used in this experiment were those we generated for the two experiments in section 4.2. There were 12 peak locations and two signal types (speech and noise). We presented each stimulus twice, which resulted in a set of 48 stimuli for which P-centres should be determined. We selected 12 phonetically or musically trained subjects who could be expected to perform adequately on the difficult rhythm adjustment task. Eleven of these subjects also participated in the previous experiments. None of the subjects reported any hearing difficulties, and none received any payment for their participation.

To determine the P-centres in our stimuli we made use of a programme that uses a rhythm adjustment task much like that described in Scott (1993) to determine the location of P-centres in all possible sorts of stimuli.[6] The programme creates 5 timing marks and produces clicks (0.5 ms, 8 kHz square wave pulses with an interval of 700 ms) on the first and the last two. The third mark does not receive a click, but a stimulus (in our case a syllable) is presented randomly within the 1400 ms between the second and fourth timing mark. The subject is asked to make the sequence sound rhythmical by moving the stimulus to the right or left within the 1400 ms interval. Movement is regulated by a series of screen "buttons" which can be clicked with the mouse to shift the stimulus. The buttons have the values: + or – 5, 10, 20, 50, 100 and 200 ms. Figure 5 illustrates one trial in the experiment.

---

[6] The programme was developed at the Institute for Perception research (IPO) in Eindhoven by Dik Hermes whom we thank for his kind permission to use it.

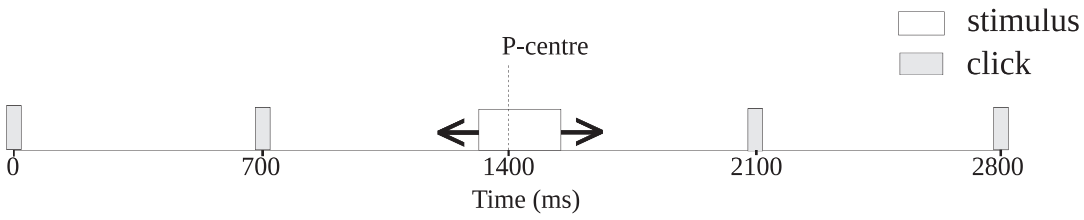


Figure 5: one trial in the P-centre experiment

The experiment was held in a sound proofed booth, and the P-centre programme ran on an Iris Indigo workstation with earphones and a mouse attached to it. After the stimulus with its accompanying sequence of clicks was presented, subjects could adjust the position of the syllable by clicking one of the buttons. After this, the whole sequence was made audible again. The subjects were allowed to change the position of the stimulus as often as they liked. The adjustments they made were cumulative, so that shifts to the right (+) or to the left (–) of more than 200 ms were made possible. When the subjects were convinced the sequence was rhythmical they pressed a key on the keyboard to store the P-centre value in a result file and initiate the next trial. Three practice stimuli were presented to familiarise the subjects with this type of experiment. To determine the location of the P-centre the programme calculated the duration from the onset of the stimulus to timing mark three (see Pompino-Marschall 1991). This is the point at which the third click would have been placed, had it been present in the series. The stimulus that is placed in the 1400 ms interval replaces the third click. So, its P-centre should be located right on the third timing mark. Therefore, the duration from the beginning of the stimulus to the third mark was stored as the value for its P-centre.

4.3.1.2 Results and discussion

First we determined the accuracy with which the subjects performed the task. Three subjects, whose correlations of the adjusted P-centres for first and second presentation of the stimuli were below $r=0.3$ ($p>.05$), had to be removed from the dataset. We then calculated the mean P-centre for each peak position for both speech and noise. These means are presented in figure 6. The P-centres are plotted expressed in terms of their distance (in ms) from the beginning of the stimulus (y-axis). For each intensity peak location (x-axis) two P-centres are plotted, one for speech and one for noise (exact figures can be found in appendix C).

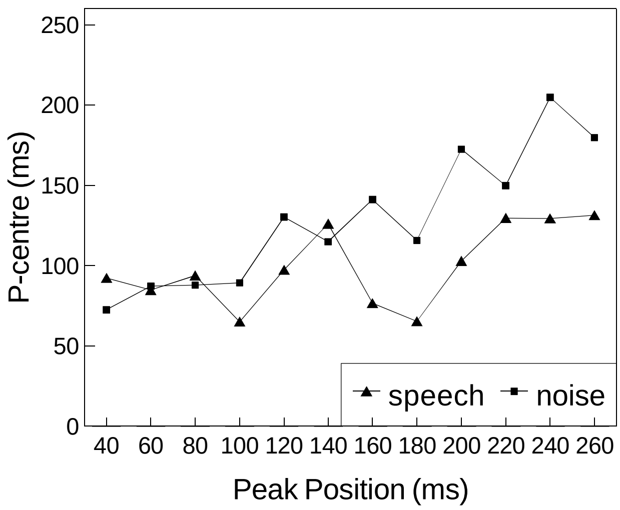


Figure 6: location of the P-centre as a function of peak position broken down by stimulus type.

Let us first look at the noise stimuli. It is clear that the P-centre moves to the right edge of the syllable, together with the intensity peak. Remarkable is the lack of P-centre movement for the intensity peak values below 100 ms. In those cases the P-centre is steady at about 80 ms, only to rise to higher values when the intensity peak moves beyond the first 100 ms of the stimulus. For peak values above 100 ms the P-centre location is dependent on intensity peak position. The unexplained alternating pattern that can be observed obscures the trend somewhat. However, the correlation coefficient between P-centre and peak position is $r=0.48$, $p<.001$. In our view, it suffices to show that the P-centres for the upper and lower end of the peak-position factor are significantly different. Since the perceived durations that we measured for these peak positions are statistically equal (cf. sections 4.2.1.3 and 4.2.2.2) a significant difference in P-centre position may lead to the conclusion that the location of the P-centre does not influence the perceived duration of noise stimuli. Indeed, a one-way anova with post hoc SNK range test shows that the P-centre values for the early peaks (40-100 ms), the middle region (120-180 ms) and the late peaks (200-260 ms) significantly differ from each other: $F(2,231)=21.1$ ($p<<.001$). Hence, we conclude that the perceived duration of noise stimuli is independent from both intensity peak and P-centre position.

The dependency relation between peak position and P-centre for the speech stimuli is far less straightforward than for the noise stimuli. If we consider the curve as a whole, we detect a global tendency for the P-centre to move to the right together with the intensity peak. Admittedly the trend is very faint, r=0.18, p<.01. What is far more striking is the shape of the curve. During the first 100 ms of the stimulus (the onset of the syllable *mam*) the location of the P-centre only marginally changes. Then, at the onset of the vowel, it moves to the right rapidly in perfect harmony with the changes in intensity peak position. After the intensity peak passes the middle of the vowel (at 150 ms) however, the P-centre shifts back to the left, only to restart its movement to the right at the onset of the final consonant (at 200 ms). These reflections of syllabic structure in the dependency between P-centre location and intensity peak position are not likely to be coincidental. It is to be expected that the other syllabic properties that determine P-centre location influence the shape of the curve. The fact that such properties are absent in noise causes the relation to be less complicated in those stimuli. In this light, the fact that the two "dips" in the curve occur at 100 and 180 ms, near the CV and VC transitions, may not be a coincidence either. The two small "peaks" in figure 4 happen to occur at exactly the same points. One could imagine one single phenomenon causing the peaks in duration perception and the troughs in P-centre location. Determining the nature of this phenomenon, though interesting, falls outside the scope of this thesis entirely. It is not our concern, since the peaks in figure 4 cannot have been caused by early P-centre occurrence. P-centre location and perceived duration are not related in our speech data, as we will see immediately below.

The odd shape of the speech curve is reflected in the correlation coefficient for P-centre with peak position (r=0.18, p<.01). It is quite low and has no real meaning if the line through the data points is not nearly straight. Other statistical measures are called for. Remember that we demonstrated the significant difference in P-centre locations of the noise stimuli for three groups only (early, middle and late peaks), a strategy that seems applicable to the speech curve as well. In this case we use 4 groups, doing more justice to the shape of the curve: early (40-80 ms), rising (100-140 ms), falling (160-200 ms) and late (220-260 ms). So, if variation in the P-centres of these groups is significant while variation in perceived duration is not, we can discard the P-centre as the duration perception trigger for speech as we have done for noise. In a one-way anova we find: F(3,229)=4.0, p<.01. The SNK range test indicates the groups with the late and the falling peaks are significantly different, as are the late and the early peaks. Just as for noise, P-centre variation is negligible if the intensity peak occurs in the first 100 ms of the signal. In any case, the large local variation that we find in the rest of the domains should be reflected in the perceived durations measured for the stimuli that belong to these peak values. Like for noise, we have found that the perceived duration of the speech stimuli does not vary. The significant fluctuations in P-centre location, combined with the fact that comparable local fluctuations cannot be found in figures 2, 3 and 4, show us that P-centre location and perceived duration are not related.[7] Therefore, we must abandon the hypothesis that the P-centre is the trigger that starts the mechanism with which we measure duration.[8]

4.3.2 General discussion

With respect to the trigger hypothesis (2) which we put forward in section 4.1 we can now draw the following conclusions. Firstly, since the position of the intensity peak does not influence the perceived duration of speech and noise stimuli, we may not assume that duration perception only starts when it is triggered by the occurrence of an intensity peak. Secondly, we find no influence of P-centre location on the perceived duration of speech and noise stimuli. Hence, the P-centre cannot trigger accurate duration perception any more than the intensity peak can. We have come close to rejecting the hypothesis that onset duration is poorly

---

[7] To push this evidence even further we calculated the F-value for perceived duration using the same four groups and the data from section 4.2.2.2 (data from the pairwise-comparison experiments could not be used since anovas are not valid in those cases). We found: $F(3,1651)=2.6$, ins. The fact that the differences between these groups are not significant justifies the conclusion that perceived duration does not vary in the domains where P-centre location varies significantly.

[8] Remember our note on C-centres in chapter 2 (footnote 9). There we cited Browman & Goldstein (1988) who found the location of the articulatory midpoint of an onset-consonant cluster (the C-centre) with respect to the VC transition to be fixed (while C-centres of coda-consonant clusters move to the right with respect to both CV and VC transition as the cluster lengthens). They put forward the possibility of a relation between onset weightlessness and invariant C-centre location. We have claimed that such a relation can only really explain the onset's weightlessness if the C-centre is perceptually relevant, and if we can find a perceptual reason for the fact that its invariant location should lead to weightlessness. With the knowledge we now have we could have proposed the C-centre as a fourth possible trigger. However, as far as we know, the C-centre cannot itself be perceived. Hence, it cannot function as a trigger. Browman & Goldstein propose the P-centre as a perceptual correlate of the C-centre. If that relationship is valid, then we know C-centres and onset weightlessness are not related, since P-centres and duration perception (weight) are not related. However, much about the relationship between P- and C-centres, and even about the true nature of these points themselves, remains unclear. Further research in these fields might uncover more evidence that may lead us back to Browman & Goldstein's suggestion. For lack of such evidence we leave the issue open for discussion and continue our search in another direction.

perceived because accurate perception of duration must be triggered by a salient syllabic phenomenon that occurs after the onset. The hypothesis may be saved though, if we should find a trigger for which a change in position is reflected in the perceived duration. We already know that this is true for the PIVOT. The steep rise in sonority or the rapid formant transitions that characterise the PIVOT may act as the trigger we are looking for. Indeed, in chapter 3 we saw that lengthening of syllabic material that comes after the PIVOT (a relative change in PIVOT position) leads to significant increases in perceived duration, while lengthening of segments before the PIVOT only marginally affects the perceived duration. This may be interpreted as evidence for the hypothesis that accurate duration perception only starts after the occurrence of the PIVOT. However, as was noted above in section 4.1, the PIVOT is inseparably linked to the transition point at which the onset ends and the nucleus starts. Therefore we can only change the PIVOT position if we change segment durations, as we have done in chapter 3. In doing so we might also have altered some other syllabic property that may just as well be the cause of the poor duration perception in onsets. In that case we could be falsely attributing the effect to the PIVOT if we take it to be the trigger for accurate duration perception (in accordance with hypothesis 2 in section 4.1). To prevent that from happening we will abandon the trigger hypothesis for now and consider some likely alternatives.

Remember that we also proposed as a possible cause of onset weightlessness the idea that something in the acoustic signal of the onset prevents our hearing mechanism from correctly perceiving duration. To support this view we might look for differences in the acoustic signals of nuclei and codas on the one hand and onsets on the other. One such difference is the shape of the intensity curve. In onsets it rises while in nuclei and codas it is generally level or falling. As a sneak preview to investigations concerning an alternative to the trigger hypothesis, which claims that certain properties of the onset jam proper duration detection (cf. section 4.1, hypothesis 1), we may provisionally conclude from the data presented in section 4.2 that it is probably not the rising intensity in the onset that is the jamming factor. In our variations of intensity peak position we implicitly altered the rise times (the later the peak, the longer the rise time). If duration perception is bad when the intensity is rising we should have found perceived duration to shorten as rise time lengthened. We did not find such shortening, and may thus discard rising intensity levels as a jamming factor for duration perception.

Other characteristics that are typical for onset signals exist. From observations made by Wever (1949) we conclude that there might be a difference in the rate of change in neural firing between onset and coda signals. Furthermore, Handel (1994) describes a difference in the onset and offset times for signals of different frequencies, which implies that in such a complicated signal as a speech sound the durations of the harmonics are equal but shifted in time with respect to each other, the lowest starting and ending the latest. These and other considerations might form the basis for new psychophysical experiments that could reveal the cause of onset weightlessness. We fear, however, that such experiments would lead us too far into the field of neurophysiology. It is not our intent to go that far astray, and we would gladly leave it to experts in that field to look for dependencies between neurological onset properties and duration perception.

Instead of venturing on the neural pathways we will stay closer to home and devote some attention to an alternative hypothesis. In the previous sections we have ignored the fact that we have not really been able to determine the causal relation between phonological weight and perceptually invariant duration. In other words, we do not know whether: 1) *onsets are weightless because duration differences in onsets are not perceived accurately*, or 2) *whether differences in onset duration are poorly perceived because onsets are weightless*. In case the reason for the asymmetry in duration perception that we have found in chapter 3 cannot be found in psychophysical experiments, invalidating 1), the reasoning in 2) might apply. Poor perception of duration variations in onsets might be caused by the fact that onsets are marked as weightless in the abstract representation of the syllable in our heads. That would mean the experiments in chapter 3 have not revealed the cause of onset weightlessness, but an effect of it. In a way, such an answer would not be very satisfactory. The question why onsets are weightless would remain unanswered. We would only have determined that onsets must indeed be weightless (since we have the phonetic evidence), and that the cause for this weightlessness must be found outside the fields of phonetics and psychophysics (possibly in psycholinguistics).

So, it might be the case that the cognitive model that represents the syllable in our heads necessarily involves weightless onsets (for whatever reason). This weightlessness could then be reflected in the poor perception of duration differences. As such, the relative perceptual invariance of onset duration would be a phonetic correlate of a phonologically determined phenomenon. Therefore, we predict that it should be unique to speech. For non-speech signals the linguistic syllabic model is irrelevant. Hence, it should not force weightlessness on the onsets of non-speech stimuli. Consequently, duration differences in onsets of non-speech stimuli would be as correctly perceived as the durations in the rest of the signal.

The underlying hypothesis that we will test in our final experiment is: *relative perceptual invariance of onset duration is caused by the fact that onsets are weightless in the abstract representation of syllables.* We will test this hypothesis by creating an artificial non-speech “syllable” of which we vary “onset”, “nucleus” and “coda” durations in exactly the same manner as we did in chapter 3. We will conduct a pairwise-comparison experiment using these stimuli. If our hypothesis is correct, no perceived duration differences between onset, nucleus and coda will be found for this artificial syllable.

4.4 Duration perception in artificial non-speech syllables

In this section we will introduce an experiment similar to that described in section 3.4. This time, though, we will need to use different stimuli, because we wish to check whether the asymmetry in the duration perception of onsets, nuclei and codas is also present in non-speech signals. We want to replicate the original pairwise-comparison experiment that was presented in chapter 3 as closely as possible. We will leave out the base duration variable, though, since we found only a small effect for base variation in the previous experiment. We will use a 300 ms base only. Not just any non-speech signal may serve to compare the outcome of this experiment with that of the speech experiment. We need a non-speech signal that is divided into three parts, and which resembles a syllable in its broader sense. The durations of the three “segments” of the non-speech syllable will then be adjusted in exactly the same manner as the speech segments in the former experiment. The task for the subjects will again be to determine stimulus duration through pairwise comparison. As noted above, the prediction is that we will find no difference between the three “subsyllabic constituents” this time, since our hypothesis states that the perceived duration phenomenon we found in chapter 3 should be unique to speech.

4.4.1 Stimuli

The non-speech stimulus that we needed was constructed from two USASI (100 Hz Hp + 320 Hz LP, both with 6 dB/octave cutoff slopes) noise parts which flanked a sawtooth signal. The frequency of the sawtooth was 100 Hz. As such, the stimulus resembled a real speech syllable in having a fricative onset and coda, and a periodic nucleus. With some imagination the signal could be perceived as Dutch *sas* ‘good humour' or *soos* 'youth club'. The three constituents of this signal were 100 ms long. The resulting reference stimulus appeared to be a perfectly symmetrical syllable, as can be seen in figure 7.

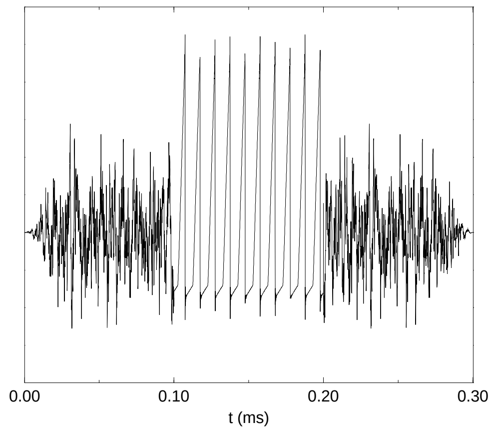


Figure 7: oscillogram of the artificial non-speech syllable.

The (linear) rise and fall times were set to 30 ms to avoid disturbing clicks. Then the comparison stimuli were created by lengthening or shortening the durations of "onset", "nucleus" and "coda" in the reference stimulus by 20, 40 or 60 ms. These duration adjustments were carried out through cutting material from the original 100 ms segments and linking the remainder to the other two pieces while avoiding brisk intensity transitions. To prevent onsets and codas in the maximally shortened cases from becoming almost inaudible we adjusted rise and fall times in the manipulated segments in a relative way such that they constituted 30% of the total segment duration, as is the case in the original. Note that the way in which we altered the duration of the speech stimuli in the experiment in section 3.4 (by manipulation of the frame durations) also leads to relative rise and fall times (30% in all cases). We thus generated the 18 comparison stimuli presented in table II (compare table V in chapter 3). Together with the 300-ms original there were 19 stimuli.

***Table II**: onset, nucleus and coda duration (in ms) for each of the 18 comparison stimuli.*

No	O	N	C	No	O	N	C	No	O	N	C
1	80	100	100	**7**	100	80	100	**13**	100	100	80
2	60	100	100	**8**	100	60	100	**14**	100	100	60
3	40	100	100	**9**	100	40	100	**15**	100	100	40
4	120	100	100	**10**	100	120	100	**16**	100	100	120
5	140	100	100	**11**	100	140	100	**17**	100	100	140
6	160	100	100	**12**	100	160	100	**18**	100	100	160

4.4.2 Method

Thirty-three Dutch subjects participated in the experiment. Their ages varied between 18 and 45, and none of them reported any hearing difficulties. The subjects were not paid for their participation. Stimuli were either pairs of the reference and a comparison stimulus or repetitions of the reference stimulus, with a silent interval of 500 ms in between. The two possible orderings reference-comparison and comparison-reference were both exploited, raising the total number of sets to 38. These sets were put in random order and presented twice to the subjects via earphones attached to an Iris Indigo workstation that produced the stimuli. The subjects had to decide which of the two members was the longer one by pressing one of two keys on the keyboard. They also had the possibility to repeat the trial by pressing the spacebar. Once a response was given, the computer automatically turned to the next trial. The subjects' responses were recorded and stored on disk.

4.4.3 Results and discussion

Like in the pairwise-comparison experiment presented in chapter 3 we calculated the percentage of subjects that marked the comparison stimulus (C) the longer one for each stage of duration manipulation. We expected the sigmoid psychometric curves to resemble the s-shape we also found in the curves given in chapter 3, figure 4 (for the same reasons that are given there). Indeed, percentages reflecting "C longer" judgements should drop as the stimulus is shortened, and they should rise as it is lengthened.

However, if duration in non-speech signals is perceived drastically different from duration in speech signals, we expect the psychometric curves to reflect that. The most straightforward results we could get in that respect would show three nearly identical curves. This would mean that duration perception does not discriminate between the three non-speech “segments”, as it does for speech. In figure 8 (and appendix C) the results of this experiment are presented. Percentages expressing the number of subjects that judged a particular stimulus to be longer than the 300 ms base are given as a function of the size of the duration manipulation (x-axis). Three curves are given, one for each “segment”. This figure shows us that there is indeed no difference between “onsets”, “nuclei” and “codas” in the duration perception of artificial non-speech syllables. Apart from a spurious data point for the stimulus in which the “coda” was lengthened to 120 ms, the three curves are nearly identical. We have no explanation for the single odd data point. Since the rest of the coda curve remains closely to the other two curves, reflecting similar behaviour for the three constituents, we seem to have no other option but to label this odd data point as an artefact of our experiment. Hence, we consider it to be of little significance with respect to our hypothesis.

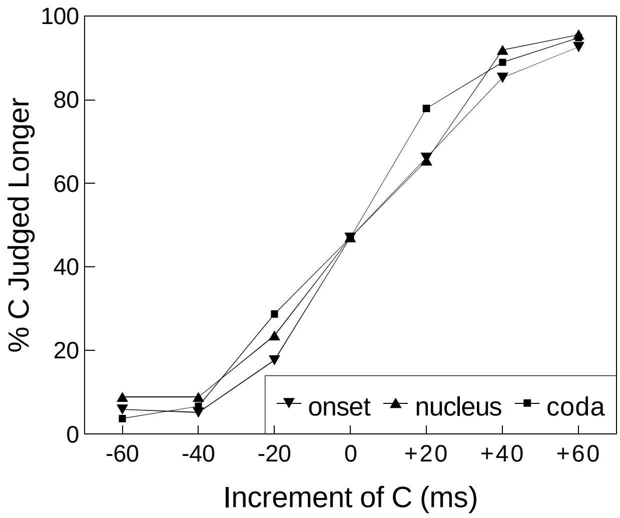


Figure 8: overall percentages of longer judged stimuli broken down by duration manipulation and segment type.

The curves in figure 8 most closely resemble the curve we found for the nucleus in the speech experiment. This resemblance is reflected in the JNDs. As we did previously, we determine JNDs by taking the abscissa with the x-axis corresponding to the points at which the subjects’ judgements are correct in 75% of the cases (hence, we look at 25% for shortening and 75% for lengthening). The JNDs we find, collapsed over lengthening and shortening and expressed as a percentage of the reference stimulus duration are 7.7%, 7.6% and 6.9% for nuclei, onsets and codas, respectively. All these percentages lie in the JND range we found for the nucleus earlier (cf. table VII, chapter 3). The JNDs are so close together that we take duration perception in “onsets”, “nuclei” and “codas” of non-speech stimuli to be identical.

A premature conclusion we might draw from these results is that there is a huge difference between speech and non-speech as far as the perception of duration variation in sub-segments of the signal is concerned. Even the small differences that we do find in the curves in figure 8 run counter to what we might expect if duration in speech and non-speech were to be perceived in the same manner. We observe that, for shortening, the onset does marginally better than the nucleus which is exactly the opposite of what we found for speech.[9] Therefore, we might assume that the perception of duration depends on the nature of the signal. Durations of all segments of an arbitrary acoustical signal are perceived in the same fashion. When the signal is a speech utterance, though, the additional value of that fact distorts the perception of duration such that onsets, nuclei and codas behave differently, as we have seen in chapter 3.

However, we must interpret the results of this experiment with some caution. We do not claim to have had the ultimate wisdom to create the perfect artificial syllable. The stimulus we used was a first approximation, but it may have been imperfect in some ways, such that it was unsuited to produce the differences in duration perception we were looking for. Perhaps the two noise bursts and the periodic signal that made up our syllable sounded too discontinuous to trigger the effect. It might be that subjects must perceive the stimulus as one continuous signal coming from one and the same source for the duration asymmetry to take effect. Moreover, there might be properties that are encoded in the speech signal that we have not simulated in the non-speech signal. These might be responsible for the durational asymmetry effect.

Further phonetic research is needed to determine whether we have not missed anything. A change in the first formant (mouth opening), for instance, was not represented in our stimulus. If this change in F1 happens to be the trigger, it is logical that we have not found a duration

---

[9] Since differences are so marginal, we do not wish to draw any firm conclusions from this observation. For the lengthening cases we will not draw any conclusions at all since the odd data point we mentioned earlier prevents such fine-grained observations.

asymmetry in non-speech. Yet, we might have easily found it if we had used a hum-sawtooth-hum artificial syllable in which the change in F1 *is* simulated. A further experiment that we might think of is one in which we use the method of sine-wave analogons to create our artificial syllable. In this method (see Parker 1988) the formants in the speech stream are replaced by sine waves of the formant's frequency, creating F0-less "speech" with discrete formant values tracing the formant tracks that could be found in the speech signal. It might be that exactly those properties of speech that are necessary to trigger duration perception differences are retained in the sine-wave analogon stimulus. A rapid change in the formants, for instance, can still be detected. Hence, a PIVOT equivalent is still present in the signal, possibly acting as a trigger for accurate duration perception.

Remember in this respect that we were not able to determine whether the PIVOT was the trigger in the first place, though we could not rule it out either. If the sine-wave analogons produce the effect, the hypothesis in which the PIVOT is the duration perception trigger comes back into the picture. As we noted in section 4.1, confirmation for the PIVOT as the trigger must come from an experiment in which the segment durations of onset and nucleus remain unaltered, which seemed impossible. An experiment that might provide some more indirect evidence, however, is the following. Suppose we could make syllables in which the location of the PIVOT is more sharply defined than in normal speech. We could try to achieve this by shortening the rise time of the sonority curve at the CV transition. Furthermore, we could make the formant transitions there more steep. Thus, the listener can locate the position of the PIVOT in the syllable more exactly. If this leads to an increase in the accuracy of duration perception after the PIVOT, we have found indirect evidence for the PIVOT-trigger hypothesis.

Unfortunately the limited time we have for this study prevents us from undertaking these experiments. They must be carried out, however, to make sure we do not falsely attribute onset weightlessness to a purely linguistic phenomenon. Only when all non-speech signals lack the asymmetry in duration perception we find for speech, and when we can confirm that the PIVOT is not a trigger in this matter, we may embrace our linguistic alternative.

For now we can only conclude that there is an indication that the perception of duration in non-speech might indeed be crucially different from the perception of duration in speech. Hence, the hypothesis stated at the beginning of this section finds some support in the results of this experiment: some abstract linguistic principle that is present in our speech centre (which is activated when we perceive speech) might cause the differences in duration perception that we detected in chapter 3. Thus, relative perceptual invariance of onset duration may be caused by the linguistic model of the syllable that is present in our heads.

If the cause of onset weightlessness is indeed linguistic, then the differences in duration perception we have found cannot form evidence for the weightlessness of the syllable's onset. Rather, in these differences we have discovered a new correlate of phonological weight. Since this correlate behaves differently in onsets on the one hand and nuclei and codas on the other, we have also found new phonetic evidence for the abstract syllable division that was presented in chapter 1. However, it would also mean that we have come full circle in our attempt to phonetically explain the onset's weightlessness. In case hypothesis 2 in section 4.3.2 is true, we can offer no reason for the fact that our internal representation of the syllable contains weightless onsets. Perhaps the reason must then be sought in psycholinguistics or even neurology.

Although we have presented the conclusion that we may draw from the previous experiment with some reservations, we may take this conclusion, along with the conclusions drawn from the experiments in the previous chapters, and speculate a little on what these conclusions mean for a phonological model of syllable weight.

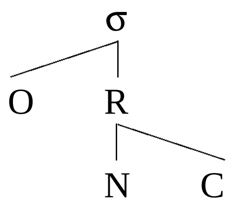
4.5 A model for syllable weight and prominence

Combining the evidence we have gathered in this chapter and the previous ones, we may be able to develop a provisional model of the relation between duration perception and weight. We depart from the following three observations:

(3)
1. perception of duration differences is better in segments that can have moras (nucleus, coda) than in segments that cannot have moras (onsets); chapter 3.
2. moras are abstract elements assigned to syllabic segments for a reason that may be purely linguistic or psychophysical in nature; chapter 4.
3. duration differences in sonorous segments are perceived as larger than equal differences in less sonorous segments; sections 3.2.2 and 3.3.2.

Any model we develop must at least cover these three observations. We must now proceed by defining the general outline of a structure relating the factors that influence perceived duration to phonological weight. For this purpose we make use of an idea that is proposed in van der Hulst & Rowicka (1997). We mentioned in chapter 1 (section 1.2.2) their suggestion that onset-rhyme syllables and moraic syllables might both exist, and be applied to different domains. Onset-rhyme syllables may be represented by the head-dependent relationships expressed in (4).

(4)



The structural differences in this figure are reflected in the accuracy of duration perception that we have studied in chapter 3. Duration differences are exaggerated in the head (N), perceived faithfully in the dependent (C), and underestimated in the specifier (O). We do not take this to be a coincidence. Yet, van der Hulst & Rowicka claim that not the onset-rhyme syllable but the moraic syllable is applicable to the prosodic domain. Hence, the moraic syllable should reflect our findings as well, or even better, drawing into the extreme the relative differences in the accuracy of duration perception that we have discovered. Since we are dealing with stress, a prosodic phenomenon, we adopt the view that the moraic syllable is the one we have to work with. Therefore, we will no longer refer to the structure in (4).

Let us now build the prosodic syllable from scratch. We start out under the assumption that every step we take must be motivated by the results of our duration experiments. As such, emphasis is put on properties of the prosodic syllable that are relevant to phonological weight. Perhaps this means we will discover in the future that our model for the prosodic syllable is oversimplified, but for now it will do. From (3) we conclude that sonority and moras are paramount. From (3.3) we deduce that the sonority difference between vowels and consonants has an impact on perceived duration, and, hence, possibly on phonological weight. We, therefore, must incorporate this difference in our model. We envisage the sonority envelope of the prototypical CVC prosodic syllable as the concatenation of three sonority building blocks, one low (call it level 1) block, one high (level 2) block, and finally another level 1 block. Note in this respect that other prosodic properties (or segmental properties that can influence sonority) we have not dealt with in this thesis may be incorporated in these blocks. We have attributed the better performance of the vowel in our duration experiments to its higher sonority, but other prosodic influences, such as pitch, may influence duration perception as well. It is even very likely that the relevant properties differ across languages. However, in the absence of evidence to the contrary we will abide by the view that sonority is the most important property in this case.

This model of sonority blocks is not complete, though. We know from our perception experiments that the elements of the prosodic syllable are not perceived on equal terms. The role of the nucleus, and often also the coda, is larger than that of the onset (cf. 3.1 above). This is where the moras come in. Either the linguistic centre in our brain, or some general psychophysical phenomenon, introduces moraic building blocks that "push up" the sonority blocks of nucleus and coda, creating the moraic syllable (3.2 above). As yet, we have no reason to believe that constituency relationships exist in this moraic syllable (opposed to the lexical syllable in (4) for which such evidence can be found, e.g. in phonotactics). In our representation we do not have to group the nucleus and the coda in one higher constituent to refer to the difference between heavy and light syllables, as we will see below. Therefore, we depart from the commonly accepted representation of the moraic syllable as it was presented in (10) in section 1.2.2. We presume that the structure dominating the moras there was necessarily put in because proponents of the moraic syllable work with one syllable model that applies to both the lexical and the post-lexical domain. We, on the contrary, can assume that no structure survives at the prosodic level, just moras. We found moras to be perceptually more relevant than sonority (cf. section 3.3.2), so we represent them with higher blocks. These blocks raise the sonority blocks of the nucleus (and optionally the coda) to new levels of perceptual relevance. Note that this raising operation is nothing more than the translation to this model of the insertion of a moraic layer in other theories (which is well motivated there, see, among others, Hayes 1995). Below we will assume that the onset sonority block remains at level 1. The mora blocks represent the next higher step of perceptual relevance, so they reside at level 2. Remember there was a difference in height between the sonority blocks of coda and nucleus. Hence, if the coda receives a mora, it reaches a level that is higher than that of the moras themselves, level 3, but the nucleus is pushed up even higher than that, to level 4. The exact shape of the prosodic syllable partly depends on whether the coda receives a mora or not. If not, the coda level remains 1. We present the two possibilities in figure 9. Note that the number of nucleus segments also determines the shape of the prosodic syllable. Long vowels are often heavy, while short vowels are light. In that case the representations in figure 9 would show two moras in the nucleus of syllables with long vowels and only one when the vowel is short.

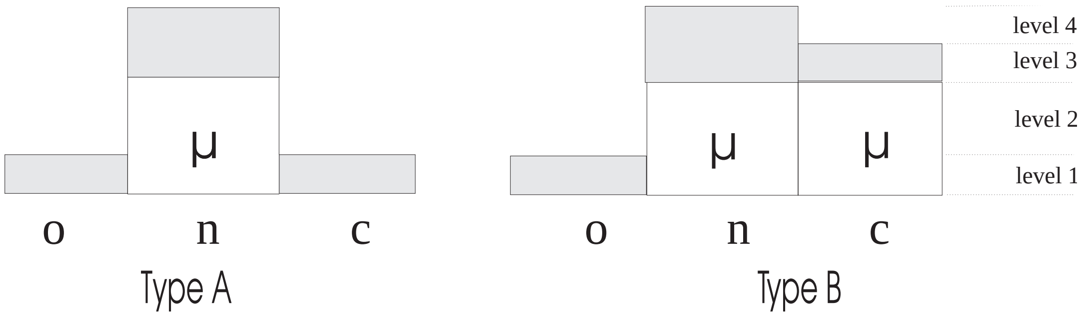


Figure 9: schematic prosodic syllables as relevant for phonological weight (dark areas indicate sonority). The difference between type A and B reflects the setting of the WBP parameter (no WBP: type A, WBP: type B).

Remember that languages in which codas are moraic use the Weight-by-Position principle (WBP, cf. chapter 1) to invoke the potential weight that codas may add. In figure 9 the WBP may be visualised as the addition of the second moraic building block to go from syllable type A to B.

We propose that, in this representation, moras and sonority (and possibly other prosodically relevant properties) conspire to determine the level of a perceptual unit of measure which we think is best described by the term *prominence*.[10] The total prominence of the syllable determines its weight. Since we only need to distinguish between heavy and light syllables we are allowed to make a binary distinction between the total prominences of these syllables. We suggest the following. In the spirit of van der Hulst (1984:69) we propose that languages employing weight to determine stress positions place a threshold somewhere on a particular level of the representation in figure 9. This threshold is used in the calculation of the minimal prominence a syllable must have to be heavy. Segments with prominences that lie above that threshold contribute to

[10] With this usage of the term prominence we deviate from Hayes (1995) who introduced prominence into metrical phonology to refer to non-moraic syllable-weight variations (see also chapter 5). When using the term prominence for the moraic level as well we also deviate slightly from its regular phonetic use (cf section 1.1.1). To be able to distinguish clearly between systems that operate at levels 1, 3 and 4, and weight systems that operate at level 2, we will refer to the former as prominence systems and to the latter as weight or QS systems, though these must be interpreted as described in this section (especially figure 9). Whenever we think clarification is necessary we will refer to the levels in figure 9. Thus, traditional QS systems are level 2 prominence systems in our view, but we will often refer to them using the traditional terminology.

weight, those of which prominences stay below that threshold do not. Usually, when two segments exceed the threshold a syllable is prominent enough to be heavy.

Most languages that use syllable weight operate at level 2: their prominence threshold lies somewhere between the lower and the upper boundary of level 2. Hence, the weighting processes in these languages are only sensitive to the very robust prominence contributions made by moraic segments: all moraic segments add to weight, while non-moraic segments do not. As was noted in chapter 1 (cf. example (8) in section 1.2.2) two types of quantity-sensitive languages are the most common. Languages like Walangama use type A syllables. Their prominence threshold lies somewhere on level 2, so every vocalic element contributes to weight. Coda consonants are not moraic, so their prominence never reaches the critical threshold and they cannot contribute to weight. Languages like Latin use type B syllables. In these languages vocalic elements exceed the threshold, as in all QS languages, and if coda consonants are present, they are raised to this level by the WBP rule. Again, if level 2 is maintained for a short period (the duration of one segment) the syllable is light, and if it is maintained for a long period (two segments) the syllable is heavy. In latin type languages, long sustainment of level 2 prominence can be achieved by long vowels (2 vocalic segments) or by a short vowel with a closing coda consonant.

These ideas are in perfect harmony with the views expressed in Beckman (1986:197) who observes that the *total intensity* (the summation of intensity over time) is a better perceptual cue for stress than either duration or intensity alone. In our model this would mean that the total surface under the prominence curve (calculated by its integral) is a measure for weight. This surface can be increased by both longer syllable duration (moras) and any segmental or prosodic property that can influence the *height* of the prominence curve (like pitch and sonority). Below we will refer to such properties as *prosodically active* and we will claim that prosodically inactive segmental properties cannot distinguish between heavy and light syllables (and, thus, influence stress placement).[11] Note that our coarse separation of heavy and light syllables on the basis of "whether syllabic prominence remains high long enough" is merely a binary abstraction of the difference between the surface integrals of heavy and light syllables.

---

[11] It is not our intention to discuss the segment internal organisation of features in this book. We must note, however, that from our point of view any theory in which it is possible to refer to segmental features that can be prosodically active as a natural class is preferable. We give van der Hulst's (1997) Radical CV-Phonology as an example.

Adoption of this model entails that we recognise some other possibilities as well. Languages may place the prominence threshold at one of the sonority levels: 1, 3 or 4. Remember, though, that these levels are subject to great variability: sonority is not the same for every possible segment, unlike moraicity. This might be of great importance to the weight phenomenon. Consider a hypothetical language of which the prominence threshold lies somewhere in level 3 (hence, uses type B syllables). In this language, some coda consonants will be sonorant enough to exceed the threshold, while others will not be. For instance, voiceless consonants may not be sonorant enough to reach the threshold: their prominence lies within level 3, but below the critical threshold. Voiced consonants are more sonorous, which means they might exceed the threshold, thus contributing to weight. Such a sifting of "heavy" and "light" codas is much more fine grained than the coarse distinctions that are made at level 2 (which is based on the presence or absence of segments). More accurate perception of prominence is needed in these cases. Therefore, languages that operate at level 3 will be scarce when compared with the group of languages like Latin and Walangama. Indeed, there are very few of these languages. An example that was already mentioned in chapter 1 is Inga (see Levinsohn 1976 and chapter 5). Note that our model predicts vowels in such languages to always exceed the prominence threshold and contribute to weight. It is not to be expected that the sonority of vowels can drop to level 3 and end up below the threshold, by definition it remains on level 4.

There must also be a level at which one can tease apart the sonority distinctions between high and low vowels. In our schematic syllable this must happen at level 4. According to Kenstowicz (1994) some languages that use this option are Kobon, Chukchee and Mari. In such languages low vowels can make a syllable heavy (or heavier) while high vowels, being less sonorous, cannot.[12]

Finally, we predict the existence of pure onset prominence systems that operate at level 1. In such systems the distinction between heavy and light syllables can be made on the basis of whether the sonority of the

---

[12] We note here that it does not really matter whether nuclei and codas are moraic in these sonority sensitive systems. The same distinctions can be made at the lower absolute prominence levels we encounter when prominence is not "boosted" by moraicity. However, for the coda case, the absence of the mora level would mean that weight distinctions are made on the level that is also relevant for onsets. Since we do not know any systems in which onsets and codas can both add weight on the basis of their prominence, we assume that nuclei and codas are moraic in all level 3 and 4 systems. This may be taken to mean that the moraic option for weight discrimination is more basic than the prominence option (as is to be expected). Languages that wish to make weight distinctions will probably always exploit the moraic option first.

onset does, or does not, exceed the prominence threshold. The fact remains that onsets cannot be moraic. Moreover, much more prominent material to form the basis of heavy/light distinctions is available in nuclei and codas. These facts make it extremely unlikely that languages need to employ the onset prominence option, but we cannot exclude it. In chapter 5 we elaborate on this possibility.

Further predictions made by our model concern the status of the other syllabic segments in cases where the threshold is located on level 1, 3 or 4:[13]

(5)
1. If the prominence threshold lies in level 4 (hence, vocalic sonority is relevant to weight), codas cannot add to weight.
2. If the prominence threshold lies in level 3 (hence, coda consonantal sonority is relevant to weight), all nuclei will add weight.
3. If the prominence threshold lies in level 1 (hence, onset consonantal sonority is relevant to weight), all nuclei will add weight (and all moraic codas).

Observe that in none of these points onsets have such a status that they contribute to weight whatever their prominence (which nuclei and codas can do). So, the mere presence of an onset can *never* make a syllable heavy. For that to happen the threshold would have to fall below the lowest level possible in figure 9. This is a consequence of the assumption that onsets cannot be moraic. It is not a prediction of our model, because it was deliberately put in, based on the observations of numerous phonologists whose descriptions of stress systems include none that would support moraic onsets. Remember that this observation was the driving force behind the formulation of our research question: *Why is the syllable onset weightless?* We have found no real answer to this question, but we did find phonetic evidence to support the phonological claim that onsets must indeed be weightless. However, we have determined that the perception of duration differences in onsets is *poor*; it is not impossible to perceive them. Hence, we cannot exclude the possibility that languages exist in which the weightlessness of the onset does not apply. However, we consider the existence of moraic onsets to be extremely unlikely, since we are still convinced that some general principle causes the non-moraicity of onsets to be a linguistic universal. On the other hand, languages in which sonorous onsets may contribute to weight while non-

[13] It is not our intention to check the validity of these predictions. We merely present them as possible testing cases for the model in figure 9.

sonorous onsets do not, may occur even according to our model. As noted above, languages for which the prominence threshold lies at the onset level (1) will be rare, but are not impossible.

Evidence to counter the views presented here has been put forward by Davis (1985) in the form of languages that seem to have moraic onsets. Other phonologists (like Everett 1988) have introduced languages with onset prominence systems that present a challenge to the way in which we look upon such prominence. After some background information on metrical phonology, our views on onset prominence systems will be elaborated upon in chapter 5. In chapter 6 we discuss two cases that are representative for Davis' moraic onset languages (Western Aranda and Alyawarra), and show that these are open to a quantity-insensitive reanalysis that does more justice to the data at hand. In the second half of chapter 6 we will present a notorious case that has first been discussed by Davis (1985), namely Mathimathi. The onset-sensitive stress system of this language would be odd in anyone's book. However, with the help of some historical and morphological evidence we show that it can be brought back into the domain of the more regular stress systems.

3 The Perception of Syllabic Duration[1]

3.1 Introduction

In the previous chapter we have seen that the temporal asymmetry between onsets and codas we are looking for cannot be found in speech production. For speech production to be the cause of the inability of the onset to contribute to syllable weight, we should have found systematic differences between actual onset and coda durations, preferably such that durations of the onset remained invariant, irrespective of the number of segments contained in that onset. The physical durations of onsets we did find are comparable to coda durations for all the consonant cluster sizes (1,2 and 3) that we included in the experiment. Hence, we may assume that speech production data do not offer a phonetic explanation for the weightlessness of the syllable onset.

An alternative explanation for the onset's weightlessness may come from speech perception. As in the production case, invariance of onset duration relative to coda duration could be the cause of the phonological difference in the ability of these constituents to add weight to a syllable. Perceptual invariance, of course, is not the same as physical invariance. Perceptually invariant segments may show differences in the acoustic signal, but these differences cannot be perceived by the listener. In chapter 2 we argued that the physical absence of duration changes in onsets would prevent these onsets from playing a role in processes that refer to quantity, the phonological counterpart of duration. The argument was: *what is not there cannot serve contrastively*. By the same line of reasoning we can claim that: *what cannot be heard cannot serve contrastively*. It is to be expected that inaudible acoustic cues do not play a role in the phonology. Hence, to pinpoint duration perception as the cause of onset weightlessness, we have to demonstrate that duration differences between various onsets are inaudible, or at least perceived less accurately than duration differences in nuclei and codas. To be honest, it is not to be expected that duration differences in onsets are categorically

[1] The results of the experiments reported on in this chapter were published as Goedemans & van Heuven (1993) and the first half of Goedemans & van Heuven (1995).

inaudible. A simple mental experiment can tell us that, if we stretch the duration of the onset of, for example, *sun* long enough, everyone will hear the duration change. Therefore, perceptual onset invariance is expected to be *relative*, if it exists at all. In the relative invariance case, duration differences in onsets will be perceptually smaller than identical differences in nuclei and codas.

In trying to find a perceptual explanation for the absence of quantity in the syllable onset we, thus, depart from the following hypothesis: *Onset weightlessness is explained through the perceptual invariance of its duration. This invariance is relative: duration differences in onsets are perceived as less sizeable than equal duration differences in nuclei and codas.* In this chapter we describe three perception experiments that were conducted to test this hypothesis. Before we start with the first (pilot) experiment, however, we briefly review some relevant results from earlier experiments on the perception of duration.

3.1.1 Some previous experiments on duration perception

As a measure of the perceptibility of acoustic differences, researchers often use the Just Noticeable Difference (JND, or Difference Limen), which indicates the smallest increase (or decrease) in a single property of any perceptible entity that can still be detected. In duration experiments JNDs will be expressed in milliseconds, indicating the minimal amount of lengthening (or shortening) needed to make the listener hear the difference between a certain sound (the reference stimulus) and a durationally manipulated version of that sound (the comparison stimulus). Several experiments were conducted in the past to determine JNDs for various types of sounds. It appears that the JND for duration is linearly related to the duration of the reference stimulus (Stott 1935), it adheres to Weber's law ($\Delta t/t$ = constant) and can be expressed as a percentage of the reference stimulus duration. This percentage depends on the auditory capacities of the subjects and the method of experimentation, but is generally found to be around 10% (Abel 1972; Fujisaki, Nakamura & Imoto 1975).

One experimental method that enables us to find these JNDs immediately springs to mind. In this method subjects are presented with sets of one constant reference stimulus and one durationally manipulated comparison stimulus. Subjects are asked whether the two stimuli are the same or different. By agreement the JND is the duration change in the comparison stimulus that is needed to make 75% of the subjects judge the two stimuli to be different. Thus, the JND is a measure for the *accuracy* with which duration is perceived. A paradigm that is one step away from this same-different method is the Two Interval - Two Alternative Forced Choice (2I-2AFC) method. In that paradigm subjects are asked to choose which of the two stimuli is the longer one. The JND is the duration change (lengthening or shortening) in the comparison stimulus that is needed to make 75% of the subjects *correctly* select this stimulus or the reference stimulus as the longer one. Since in the 2I-2AFC method subjects have the extra task to label the direction of the duration difference they hear, this method eliminates the possibility that “different” judgements made by the subjects for the wrong reason (e.g. the reference stimulus was heard as shorter than the comparison stimulus while it was longer) are counted as “correct” in the computation of the JND. Abel (1972) used this 2I-2AFC method to determine the JNDs for 1000 Hz sinusoids and several types of noise burst. Her reference durations varied between 1 and 1000 ms. As noted above, she found JNDs of about 10%. Fujisaki et al. (1975) compared 500 Hz tones with two types of white noise burst (narrow and broad band noise), both varying in duration around a 100 ms base. They found JNDs of 9.6 ms and 9.1 ms for the sinusoids and the wide band noise, respectively. The JND for narrow band noise was 6.7 ms.

A second method to test the listener’s sensitivity to duration differences is to create several reference stimuli of different durations, and present these one at a time in a pair with an adjustable comparison stimulus. In a series of repetitions, the subject then has to adjust the duration of the comparison signal until he is convinced that reference and comparison stimulus are perceptually equal. The main advantage of this method is that the subject’s adjustment for the comparison signal directly reflects the actual perceived duration of the reference signal, which we cannot obtain when using the 2I-2AFC method. In other words, this method allows us to check the *fidelity* (or faithfulness) with which the listener perceives and reproduces the duration of the reference stimulus. The duration can be underestimated (the comparison signal is adjusted to a duration shorter than that of the reference stimulus), reproduced faithfully, or overestimated. Since we are interested in the perception of duration *differences*, the fidelity with which one isolated stimulus would be perceived is of little concern to us. What does interest us greatly, however, is the fidelity with which a duration *change* in, say onsets, nuclei or codas is perceived (i.e. whether the change is underestimated, overestimated or reproduced faithfully). The over- or underestimation factors we may find are *not* the JNDs for duration perception, but one can imagine the two are closely connected. Suppose a change in the duration of a certain stimulus were perceptually underestimated. Then the perceived duration for this stimulus would be shorter than its actual duration. This would mean that the actual duration change needed to reach the smallest duration step that can be perceived is *larger* than the duration change that would be needed when the actual stimulus duration were perceived correctly (which, in its turn, would be larger than the change needed in case the duration were overestimated). In other words, it is more difficult to reach the JND if the duration of the stimuli is underestimated, and less difficult if the duration is overestimated. Hence, we expect the JNDs and the over- and underestimation factors to be highly correlated.

Unfortunately, a drawback of the Adjustment Paradigm described here is that the calculation of JNDs proceeds along less well established lines than in the 2I-2AFC experiments, if they are calculated at all. Burghardt (1973) used the adjustment method to test the sensitivity of the human ear to duration for tones of different frequencies. He found that durations of tones between 1000 and 5000 Hz in the reference stimulus are systematically overestimated (the comparison stimulus is adjusted to a duration longer than the reference stimulus) while reference durations of tones lower than 1000 Hz were systematically underestimated. Van Heuven & van den Broecke (1982) use the adjustment method to show that there is a perceptual duration difference between signals with abrupt (10 ms) and gradual (50 ms) amplitude offsets for various stimuli with durations between 100 - 450 ms. They present overall JNDs of 11.4% (abrupt) and 9.6% (gradual), defined as the standard deviation of the adjustments (cf. Cardozo 1965). More importantly, they included one speech signal in their experiment. The JNDs they report for duration perception of a synthesized Dutch vowel /a/ are almost equal to the averages mentioned above (11.4% and 9.3% for 10 and 50 ms fall-times, respectively).

Other experiments on duration perception of speech can be found in the literature. Fujisaki et al. (1975) report JNDs for Japanese synthetic speech of 7% for vowels and 10% for fricatives, using the 2I-2AFC method. Klatt & Cooper (1975) describe a magnitude estimation task in which subjects had to estimate the duration of the vowel /iː/ and the fricative /ʃ/ in context using the integers 1-9. They find fairly sizeable JNDs of about 15% for /iː/ and about 30% for /ʃ/. More studies in which the perceptual duration of consonants is measured are difficult to find, and as far as we know, studies in which duration perception in onset and coda consonants is compared are absent altogether. In the pilot experiment described in the next section an attempt at such a comparison is made. Keeping in mind our initial hypothesis, we are interested in how much of a duration difference in onsets, nuclei and codas will be reflected in the perceived duration of these segments. We wish to find the fidelity with which duration differences are perceived. Therefore, the Adjustment Paradigm that was described above will be used to systematically compare duration perception in onsets, nuclei and codas. Judging from the results found in the literature we must present the speech stimuli in isolation, not in context. The perception of duration differences is better for isolated sounds than for sounds in connected speech. This may be due to the fact that the beginning and the end of the stimulus are not well defined in connected speech. It may also be due to masking, or to influence of the information that is presented in the sentence, but finding an explanation for the observed difference is not the issue of this study. We merely note that the differences we are looking for might be very subtle and may be found only under optimal conditions. Hence, we use isolated speech sounds in our experiments.

3.2 A pilot experiment

To learn something about duration perception in the three subsyllabic constituents we could take a syllable of which onset, nucleus and coda have equal durations, and systematically vary those durations. If we then present pairs of stimuli to the subjects and ask them which is the longer of the two (the 2I-2AFC method described above), we expect them to choose correctly for most pairs when these pairs contain a fixed reference syllable and a variable comparison syllable of which the *vowel* is lengthened or shortened. Pairs in which the comparison syllable shows duration changes in the *coda* should pose no problem to the subjects either, but as far as the *onset* is concerned we expect the subjects to perform poorly. This would indicate that people are more or less “deaf” to the duration differences in the onset, but it would tell us nothing about the actual duration they perceive. Nor would it precisely tell us what the ratios are between the actual duration difference and the perceived duration difference, and whether subjects can reliably estimate the duration of syllables in the first place. In other words, we learn nothing about the fidelity of duration perception in the subsyllabic constituents. Therefore we opted for the method with the adjustable comparison signal (cf. section 3.1.1). The basic idea is the same as in the 2I-2AFC method: we create a *base syllable* with equal onset, coda and nucleus durations and systematically vary the durations of these three constituents, creating a set of reference syllables. We then present subjects with these reference stimuli, each one in a pair together with a comparison signal. Subjects must adjust the second signal until it is perceptually equal in duration to the reference syllable. We hypothesize that the duration modifications made in nuclei and codas will be reflected in the adjustments the subjects make in the comparison signal. For stimuli with manipulated onset durations, however, we expect the adjustments of the comparison signal to be closely comparable to the adjusted duration we find for the base syllable (which is, of course, included in the set of reference stimuli), reflecting only small perceptual differences between this base syllable and syllables in which onsets are durationally modified. The absence of a perceptual difference in this case would seem to indicate “deafness” for duration differences in onsets, which we would take as evidence for their weightlessness.

3.2.1 Stimuli and method

The stimuli in the present experiment deviate from those used in chapter 2 in one major respect. In the production experiments we used words in which onsets and codas were varied in duration by the addition of segments. We believe that we cannot adopt this strategy here, because we do not know the psychophysical and psychological effects that these extra segments have on duration perception, so we cannot experimentally control them. It will be difficult enough to interpret the results of an adjustment experiment with a non-homogeneous reference stimulus like a syllable (which is not a steady-state sound, as opposed to most of the stimuli in the experiments described above). So, we must at least keep the segmental factor as stable as possible. To that extent we used one reference syllable with a monosegmental onset and coda and adjusted the durations of the three subsyllabic constituents, not by adding segments, but by manipulating the durations of the segments that were already there.

The basic reference stimulus was the Dutch word *sas* /sɑs/ ‘good humour’, which was synthesised from diphones (LPC 10 : 5 formants/ bandwidths, 10 ms frame duration, speaker HZ, see van Bezooijen & Pols 1993 for intelligibility data). To facilitate fair comparison between onset and coda we constructed it such that the second half of the stimulus was a mirrored copy of the first half. By deleting two selected frames from the steady state portion of the vowel, its duration was made equal to that of the onset and coda consonant: 100 ms each. We expect the duration JND of the resulting 300-ms syllable to be around 10%, as in the experiments described above. Therefore, we used duration increments/ decrements of 30 ms. Six additional reference syllables were derived from the base stimulus by lengthening or shortening each of the three segments by 30 ms. Lengthening was achieved by copying of 3 non-adjacent 10 ms frames from the centre of the segment, shortening by deletion of the same frames. For all duration manipulations of the base syllable we used the LVS speech processing software for VAX/VMS (cf. Vogten 1984). The manipulated LPC-parameter files were converted to sampled data files using LPC synthesis. The resulting set of 7 reference stimuli is specified in table I.

***Table I**: stimulus types with durations (in ms) of their constituents (manipulated segments are in boldface).*

Stimulus	Onset	Nucleus	Coda	Total
1	100	100	100	300
2	**70**	100	100	270
3	**130**	100	100	330
4	100	**70**	100	270
5	100	**130**	100	330
6	100	100	**70**	270
7	100	100	**130**	330

The rise time of the onset /s/ and the decay time of the coda were set at 40 ms, to avoid audible clicks.[2] Since we wanted subjects to estimate the duration of the entire reference syllable we did not want a comparison signal that encouraged them to attend to the vowel only. Because the /s/'s at the beginning and end of the syllable (which should definitely be taken into consideration when determining the duration) are noise-like, we opted for a noise burst (white noise), rather than some periodic sound, as the comparison stimulus. The intensity of this signal was adjusted until its loudness was perceptually equal to that of the reference stimulus, as judged by a panel of three professional phoneticians.

Subjects were 24 native speakers of Dutch, with ages between 20 and 40, and without any self-reported hearing deficiencies. None received any payment for their participation. The experiment was held in a sound proofed booth with soft panelling attached to ceiling and walls to avoid reverberation. In the booth we placed an instruction sheet and a computer terminal (Visual 603 emulating a Tektronix 4010/4014 for the graphics) with a mouse and keyboard connected to it. The subjects were presented with pairs of reference and comparison stimuli (henceforth, Sr and Sc) according to the scheme in figure 1. The pairs were presented in one of three random orders (8 subjects per order).

---

[2] This 40-ms offset is almost as long as the 50 ms offset for which van Heuven & van den Broecke (1982) report appreciably lower JNDs than for 10 ms offsets.

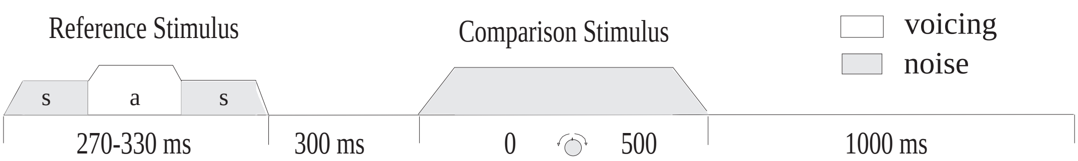


Figure 1: stimuli and pauses in one trial of the pilot perception experiment.

After Sc was made audible, a vertical crosshair appeared on the left hand side of the computer screen. Subjects were able to influence the duration of Sc, which was set at 0 ms for every new stimulus pair, by manipulating the crosshair with the mouse. Moving the crosshair to the right resulted in a longer duration of Sc, and moving it to the left resulted in shortening of Sc. The maximal duration of the noise burst was 500 ms. After this adjustment stage Sr was again made audible with a 300 ms pause following it. Then the adjusted version of Sc was presented. The stimulus pairs were repeated with 1000 ms silent intervals. During these intervals, subjects could make their duration adjustments to Sc, the effect of which was always presented during the next repetition of the pair. The subjects were allowed to adjust Sc as many times as they liked. Once they were satisfied that they had obtained the best possible match, the final adjusted duration was stored in a data file (by pressing the enter key on the keyboard) and the next trial was initiated.

3.2.2 Results and discussion

Not all subjects were able to perform their task with reasonable consistency. We considered it necessary to only look at the onset scores of subjects that were able to reliably perceive overall durations. By only selecting the subjects that perform consistently with respect to the total duration of the signal we hope to get a clear picture for the onset, at the same time removing subjects that go as far as an inverse correlation (they shorten when they should lengthen, and vice versa). Therefore, subjects whose adjusted durations correlated below r=.50 with the corresponding reference stimuli were eliminated from the data set. The onset cases were not taken into consideration in this selection criterion because: a. our hypothesis is centred around the onset scores, which are predicted to be low, and b. we wished to avoid manipulation of the data such that our conclusions with respect to the onset would crucially depend on the selection of the subjects. Seventeen subjects met the selection criterion. The overall means of the noise durations determined by these subjects were calculated for each stimulus. These means are given in table II (see also figure 1 in appendix B).

Table II: *mean adjusted durations of Sc for each of the 7 variations in the temporal structure of Sr. Stimulus numbers in parentheses refer to table I.*

Duration change	Manipulated Constituent Onset	Nucleus	Coda	Duration of Sr
30	235 (2)	195 (4)	202 (6)	270
none (base)	237 (1)	237 (1)	237 (1)	300
+30	251 (3)	313 (5)	260 (7)	330

In this table we see the mean durations to which the subjects adjusted Sc (middle three columns) for each of the seven syllables in table I. The first syllable in table I appears thee times behind ‘none (base)’, because it represents the situation in which no duration is changed for onset, nucleus and coda. For ease of comparison, stimulus numbers of the syllables in table I are repeated after the corresponding values for Sc adjustment. The actual durations of the reference stimuli, which we would expect to reappear in the three medial columns if the subjects could reproduce Sr perfectly, are given in the last column.

We observe, first of all, that Sc is adjusted to a duration much shorter than Sr across the board (the error is about 20% in all cases). We can probably not attribute this effect to the time-order error (TOE), by which the duration of the second member of a pair of stimuli is underestimated (Woodrow 1951). Underestimation of the comparison stimulus should lead to its *over*adjustment, not to its *under*adjustment (if the second member of a pair is judged to be shorter than it actually is, it must be adjusted to a longer duration to sound as long as the first member). What does cause Sc to be adjusted to such short durations remains a mystery.[3]

[3] A possible explanation might lie in a difference in the perceptual durations of noise and speech. If 300 ms of noise is perceptually longer than 300 ms of speech, then Sc must be adjusted to a shorter duration to be perceived as equal in duration to the speech in Sr, which would explain the underadjustments. Alternatively, the intensity envelope of the comparison signal might have caused the difference. We used a noise stimulus with steep offset to avoid masking effects that might interfere with our duration measurements. This abrupt offset might have resulted in a longer perceptual duration of the comparison stimulus.

In any case, it is irrelevant to our purposes. We are looking for a *relative* difference in the perceived duration of durationally different onsets with respect to such differences in nuclei and codas. The exact duration value that we find for the base reference stimulus, around which these differences centre, is of lesser importance.

If we look at table II we find the relative effect we are looking for as a difference in fidelity. The 30 ms duration manipulations in the nucleus are reproduced larger than life in the adjusted durations. In the shortening case (4) the deviation from the base adjustment (1) is −42 ms, while lengthening the vowel in the base (as in 5) leads to an adjustment of Sc that is 76 ms longer than the adjustment for the base. The adjustment differences are 59 ms on average, so, for the nucleus the duration manipulations are perceptually overestimated by a factor 2. Duration variations in the coda are more or less faithfully reflected in the adjustments of the comparison stimuli: 35 ms for a shortening of 30 ms in Sr (6), and 23 ms for a deviation of +30 ms in (7) (29 ms on average). A change in the duration of the onset, however, has hardly any effect at all: only 2 ms adjustment difference with respect to the base for a shortening of 30 ms (2), and 14 ms for a 30 ms lengthening (3). The average of 8 ms means an underestimation factor of no less than 4 for duration changes in onsets. The effects of target-segment position on mean over/underestimation of syllable duration prove to be significant in a t-test: $t(16)=-2.16$ ($p=.017$, pairwise, one-tailed) for onset versus coda, and $t(16)=-2.29$ ($p=.012$, pairwise, one-tailed) for coda versus nucleus.

Since we needed to find exact differences in perceived duration values, we used the adjustment paradigm and suppressed the need to know the (obviously related) JNDs. But we can roughly calculate them, if we follow the suggestion made by Cardozo (1965) and take the standard deviation of the adjustments as a measure for JND. We expect to find large JNDs for onsets and relatively small ones for codas and nuclei. This difference in JNDs would indicate that a relatively sizeable duration change is needed in onsets to make the subjects perceive it, while smaller changes suffice for nuclei and codas. These predictions are not borne out by the facts. In table III the JNDs are presented as a percentage of Sr.

***Table III**: JNDs for onsets, nuclei and codas by duration manipulation, calculated from the sd's of the adjustments and expressed as a percentage of the reference duration.*

Duration change	Manipulated Constituent Onset	Nucleus	Coda	Duration of Sr
–30	23	20	21	270 ms
none (base)	17	17	17	300 ms
+30	14	21	18	330 ms

It appears that the JNDs for the shortening cases are comparable for the three constituents. If anything, the JND for onsets in the lengthening case is even smaller than the one for codas, which in its turn is smaller than the JND for nuclei. Hence, the durational asymmetry between the three subsyllabic constituents that we find when we compare the fidelity with which duration differences are perceived is not reflected in the accuracy of duration perception (the JNDs). Furthermore, the JNDs are larger than the 10% that we expected to find on the basis of what was reported in the literature, though they are similar to the JNDs that Klatt & Cooper (1975) report for speech.[4] Though the JNDs in table III do not support our hypothesis, we will not draw any firm conclusions from the results found in this pilot experiment. Further discussion on the relation between the JNDs, fidelity and weight in general, and the values we should expect to find for these JNDs, will be postponed until section 3.3.2.

The conclusion we can draw from the pilot experiment is that the results presented in table II prove to be a first indication that the weightlessness of the syllable onset may be reflected in an asymmetry in speech perception. Duration variations in onsets are perceptually less salient than such variations in nuclei and codas. It seems that the hypothesis that was presented at the beginning of this chapter holds true. Duration differences in onsets are relatively poorly perceived (which comes close to perceptual invariance of onset duration). Hence, if we accept the reasoning in section 3.1, we must conclude that the onset is the least likely candidate for a phonological weight opposition. We may even

[4] Perhaps a fairer way to compute the JNDs might be to look at the individual adjustments of the subjects with respect to their individual adjustment for the base stimulus. In that way we would cancel the effect of the variable adjustments for the base on the standard deviations. Doing so, we find JNDs of about 8% for onsets, 11% for codas and 15% for nuclei. These JNDs approximate the reported 10%, but they still run counter to the expectation that the JND (the error margin) for onsets should be the largest due to the poor fidelity with which duration differences in onsets are perceived.

go as far as assuming that the perception of duration differences in onsets is so poor that it is ignored completely by the phonology. It seems logical that a phonological weight unit (mora) cannot be assigned to a constituent in which the phonetic correlate of weight is absent to a large degree. As such, onset weightlessness is the absolute phonological translation of a relative phonetic difference in duration perception.

To all intents and purposes, we must from now on assume that *faithfulness in the perception of duration differences* is one important phonetic correlate of weight. We have taken the relative absence of this correlate in onsets to be evidence for its weightlessness. However, both nuclei and codas receive the same moraic weight unit in phonological theories while our experiments show a difference in duration perception between nuclei and codas. This reveals to us that there must be a second factor at play. It is often proposed that sonority also plays a role in the determination of phonological weight.[5] The fact that subjects are most sensitive to duration variations in nuclei might be explained by the relatively high sonority of the vocalic element that fills it. It also explains the existence of many quantity-sensitive languages in which only the opposition between long and short vowels in the nucleus is used to determine syllable weight (while codas are left out of the equation). We assume, therefore, that this experiment provides indirect evidence for the secondary role of sonority in the syllable weight divisions that we find in quantity-sensitive languages. In the general discussion at the end of chapter 4 we will come back to this issue. Note that sonority cannot be used to explain the difference in duration perception that we find between onsets and codas. Both segments are equally sonorant in our experiment.

3.2.3 Shortcomings

It may be noted that we presented the confirmation of the initial hypothesis with some reservation. The caution with which we present this conclusion is fed by the realisation that what was presented here was only a pilot experiment, with its inevitable shortcomings. Firstly, the number of subjects (17) we used was relatively small. To confirm the provisional conclusion we have drawn with respect to the fidelity of duration perception it would be wise to repeat the pilot experiment with a larger number of subjects.

Moreover, in the pilot experiment there was no subject internal procedure to judge the consistency with which the subjects performed the

---

[5] As evidenced by languages in which sonorant codas add weight to a syllable while obstruents do not (cf. chapter 5).

adjustment task. Repetitions of the stimuli needed to check whether subjects make similar adjustments for the same stimulus were not included in the experiment. This meant that we had to use an ad hoc method to judge the subjects' consistency (individual correlations between Sc and Sr for the stimuli with altered nuclei and codas). Though this ad hoc method was neutral with respect to onset adjustments, we would rather use selection criteria that are completely independent from all the relevant variables used in the experiment.

Finally, as noted by Klatt & Cooper (1975), duration perception of final fricatives is not reliable. Klatt & Cooper report a very large JND for utterance final fricatives. They claim that this indicates the inability of subjects to determine the frication offset in such final fricatives. In a new experiment we would have to verify whether our usage of /s/ as the segment filling both onset and coda influenced the results. We can do this by comparing a word with an /s/ coda to a word with another coda. A good candidate might be *mam* 'mother' in which the coda is also easy to lengthen and shorten. We could also use this second stimulus to see if there is any difference in duration perception between an /s/ coda and the more sonorous /m/ coda.

In the next section we describe a larger experiment in which we try to avoid the drawbacks of the pilot experiment. Through this larger experiment we hope to confirm the results found in the pilot, and give them a more solid basis.

3.3 Duration perception in an adjustment task

Recapitulating the shortcomings of the pilot study that were listed above, we sum up the main differences between the pilot experiment and the second perception experiment and state the reason for these differences.

1. A larger number of subjects is used to enlarge the data set so that we may find more solid support for our initial hypothesis.

2. All subjects were presented with each stimulus twice. In this way, subjects that are unable to perform the adjustment task consistently can be eliminated on the basis of a comparison between the two adjustments they make for each stimulus type. In the pilot, subjects were eliminated on the basis of an ad hoc method. We do not expect the results to be influenced by this difference, but hope to give the present experiment more weight by using a more conventional selection criterion.

3. A second stimulus word (*mam*) is added so that the set of reference stimuli is doubled. The original reason for the introduction of this second word was to compare the duration perception of coda /s/'s (as in *sas*)

with duration perception in another coda segment. Such comparison is needed to check whether the poor duration perception that Klatt & Cooper (1975) report for final /s/'s, has influenced our results. A second use for the *mam* stimuli is found if we consider the fact that duration differences in more sonorant segments (vowels) were more salient than identical duration differences in in less sonorant segments (coda consonants). The difference in sonority between the two consonant types used in the new stimuli (nasal /m/ is more sonorant than fricative /s/) allows us to check whether sonority differences also lead to fidelity differences between consonants (and not just in vowels with respect to consonants). If we find a difference in the duration perception of sonorant codas with respect to non-sonorant codas we might explain the existence of languages in which these sonorant codas may add to syllable weight while non-sonorant codas in the same language cannot. Again we expect the fidelity to be higher for sonorant coda consonants than for sonorant onsets.

Apart from the expectations with respect to the difference between adjusted durations for the codas of *mam* and *sas* we expect to find the same differences in fidelity for the three subsyllabic constituents that we found in the pilot study.[6] If onset weightlessness is universal, and duration perception is its phonetic correlate, we should be able to find a duration perception asymmetry between onsets and codas a second time, and moreover, for two different words. We also expect Sc to be adjusted to longer durations for the *mam* stimuli than for the *sas* stimuli. The higher fidelity for the sonorous coda, together with a higher fidelity for the sonorous onset, which we expect to be marginal, should result in longer adjusted durations. Finally, the overall underadjustment that we found in the pilot study is expected to occur again here, at least for the *sas* stimuli, since nothing crucial in stimuli and experimental design has been changed.

3.3.1 Stimuli and procedure

The paradigm that was used for this experiment was exactly the same as that used in the pilot study. The two base reference syllables were created in the fashion described in section 3.2.1. We used computer-generated

[6] We refrain from predictions with respect to the JNDs. On the basis of the JND results in the pilot experiment we might expect no differences in the accuracy of duration perception for onsets, nuclei and codas. However, we will, of course, verify whether the JNDs we will find in the present experiment provide support for our initial hypothesis with respect to the JNDs (low fidelity means poor accuracy, means high JNDs).

diphone synthesis files of the Dutch words *sas* ‘good humour’ and *mam* ‘mother’. In the LPC-parameter files thus created we adjusted the frame durations until onset, nucleus and coda all had a base duration of exactly 100 ms. Rise and fall time were again set to 40 ms to avoid audible clicks. The three subsyllabic constituents were systematically varied through further adjustments of the frame duration which resulted in altered segment durations of 70 or 130 ms (the net result being a lengthening or shortening by 30 ms). The two sets of reference stimuli that were created both had the variable durations that are specified in table I. The LPC-parameter files were then converted to sampled data files through LPC synthesis. For the comparison signals we chose a burst of white noise for the *sas* case (as in the pilot) and a sawtooth wave (120 Hz) that sounded somewhat nasally for the *mam* case. We expected similar relations to hold between the sawtooth and the /m/ as between the white noise and the /s/, which should encourage the subjects to estimate the duration of the entire reference syllable and not only the vowel. Intensity of the comparison signals was again adjusted to achieve perceptual equality with respect to the loudness of the reference syllables.

Reference and comparison stimulus pairs were put in four random orders per stimulus type (*mam* or *sas*). A hundred and thirty-two paid subjects participated in the experiment. None of them reported any hearing difficulties. Seventy-two subjects were presented with the *mam* stimuli and 60 subjects worked with the *sas* stimuli. Hence, each of the random order files was presented to 18 subjects in the former case and to 15 subjects in the latter case. The sound-proofed booth, experimental equipment and set-up were identical to those used in the pilot study. Subjects were again presented with reference and comparison stimuli according to the scheme in figure 1. Adjustments were made in the same manner as in the pilot, during the 1000 ms silent interval. For each subject the resulting adjustments were stored on disk.

3.3.2 Results and discussion

Again it appeared that some subjects had difficulties in performing the adjustment task. Subjects whose adjusted durations in the first and second presentation of the same reference stimulus correlated negatively (excluding 70 and 130 ms onsets) were eliminated from the data set.[7] For the *mam* data 58 subjects passed the selection criterion, while 55 subjects

---

[7] The selection criterion is not based on the r=0.5 we used earlier because we now judge subject consistency on the basis of repeated measurements instead of comparison with goal values.

did so for the *sas* data. In figure 2 we see the mean adjusted durations of Sc (vertical axis) for all stimulus syllables, broken down by duration manipulation (shortened [−30 ms], base [0 ms], lengthened [+30 ms]), durationally manipulated segment (onset, nucleus, coda) and stimulus type (*sas* left, *mam* right). In each panel the intersection of the three lines denotes the adjusted duration for the base Sr of 300 ms. The other data points indicate the effect that the duration manipulations in the onset, nucleus and coda of Sr have on the adjusted durations of Sc (compare figure 1 in appendix B). The further removed from the Sc value for the base Sr, the greater the effect (and the steeper the slope of the line connecting the data points of each constituent). Tables with exact durations are given in appendix B.

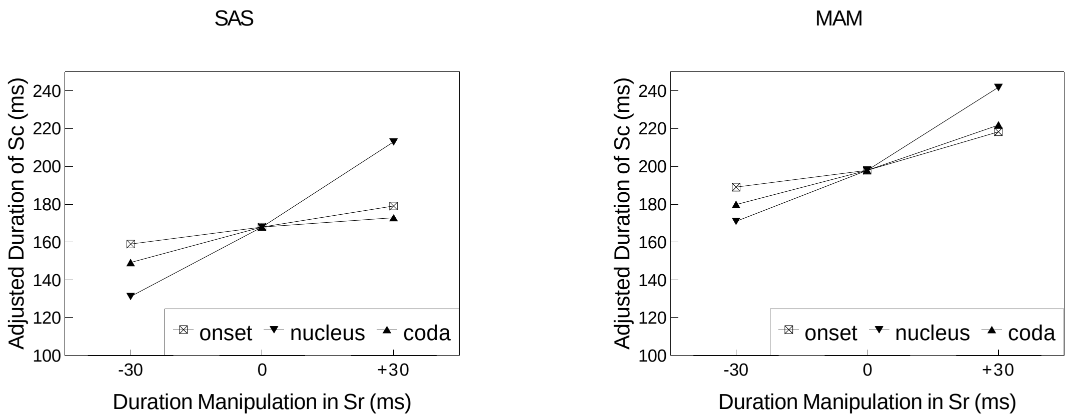


Figure 2: mean adjusted durations of Sc for stimulus syllables with lengthened and shortened onset, nucleus and coda for *sas* (left panel) and *mam* (right) stimuli.

For the *mam* data we see that the duration to which Sc is adjusted when the Sr is the unaltered base syllable (the one that has a 100 ms onset, nucleus and coda) is about 200 ms. Shortening or lengthening the onset by 30 ms should have less of an effect on the adjusted duration than such manipulations in either coda or nucleus. So, the data points for the reference stimuli in which onsets are lengthened or shortened should be closer to the “neutral” 200 ms line than the data points for stimuli in which nucleus and coda durations have been manipulated. Furthermore, on the basis of what we found in the pilot experiment we expect adjusted durations for the nucleus to be exaggerated with respect to those for the coda. Hence, we expect the line connecting the data points for the onset to be less steep than the other two, while the line connecting the nucleus points should be steeper than the coda line. We can see this happening in the right hand panel, though the difference between onset and coda in the lengthening case is very small. In the left hand panel, for *sas*, the expected divergence is only found in the shortening case [−30], for lengthening the nucleus behaves as we expected, but the difference between onsets and codas runs counter to our expectations. It might well be that this is caused by the inaccurate perception of final /s/'s that Klatt & Cooper (1975) report (cf. section 3.1.1).

If we concentrate on the right-hand panel for the moment we see that alteration of the vowel duration in *mam* clearly results in consistent adjustments of Sc duration. Adjustments averaged over lengthening and shortening are 35 ms, and the three groups; shortened, lengthened, and base are significantly different from each other with respect to adjusted duration of Sc: $F(2,345)=40.8$, $p<.05$ (one-way anova with post hoc SNK test). Sc adjustments for the manipulated coda stimuli are 21 ms on average. Again these adjustments prove significantly different for the three duration manipulation groups in a one-way anova with post hoc SNK test: $F(2,345)=16.2$, $p<.05$. The stimuli in which the onset is manipulated do much worse. Average duration adjustments of Sc are 15 ms; the one-way anova shows a significant difference in the groups: $F(2,345)=7.7$, $p<.05$ which the SNK test attributes to a difference between the mean adjusted duration for lengthening on the one hand and those for the base and shortening on the other. The fact that lengthening of the onset triggers a significant adjustment is unexpected, it should behave like shortening for which the adjustment is not significantly different from the adjustment for the base. We may conclude from this figure that the *mam* data replicate the results found in the pilot when we consider only the shortening cases. For lengthening we find that the effect of a 30 ms increase is significant for onsets as well as for codas (though both effects are again smaller than the effect we find for nuclei), and the difference between the adjustments for onsets and codas is negligible. The overall effects in terms of over- and underadjustment with respect to the actual duration changes in the reference stimuli are less dramatic than those found in the pilot. For onsets we find an underestimation factor of 2 (changes of 30 ms trigger 15 ms adjustments on average), for codas the underestimation factor is 1.5 and for nuclei there is a marginal tendency towards overestimation (compared to 4× underestimation for onsets, correct response for codas, and 2× overestimation for nuclei in the pilot).

Adjustments of the comparison signal also follow the duration changes in the nucleus of *sas* quite closely (left panel). The average adjustment is 40 ms. As for the *mam* data, the shortened, base and lengthened groups are significantly different from each other in a one-way anova with post hoc SNK test. Adjusted duration by duration manipulation of the vowel in the reference syllable: $F(2,321)=50.3$, $p<.05$. For codas we find a difference with the *mam* data. The average adjusted duration is only 14 ms, and the SNK test reveals that only the difference between the mean for shortening and the means for base and lengthening together is significant: $F(2,321)=5.4$, $p<.05$. The mean adjusted duration for onsets is only 10 ms. The effect of onset duration manipulation on adjusted duration is not significant: F(2,321)=3.1,ins, though the absolute adjusted duration for lengthened onsets is even longer than that found for codas. Again we must conclude that the effects found in the pilot are only replicated in the shortening cases. In this case the differences for lengthening between nuclei on the one hand and onsets and codas on the other are clearly visible. Lengthening the nucleus by 30 ms leads to a significant adjustment in duration of Sc while adjustments for 30 ms lengthened onsets and codas remain insignificant (as opposed to significant adjustments for the *mam* data). Over- and underestimation factors are more like those found for *mam* above than those found in the pilot. Underestimation factors are 3 for onsets and 2 for codas while we find an overestimation factor of 1.33 for nuclei.

When we look at overall underadjustments (it is best to look at the adjusted duration for the base in this respect) we find a striking difference between the *mam* and *sas* data. In all cases Sc is adjusted to a duration that is shorter than the duration of Sr. For the *mam* base syllable of 300 ms the adjusted duration is about 200 ms, a 33% underadjustment (opposed to only 20% in the pilot). The adjusted durations for the *sas* base syllable are shorter still; 167 ms on average, a 45% underadjustment. The difference between *sas* and *mam* in this respect was predicted (and proves to be significant in a one-way anova: $F(1,1466)=53.7, p<.001$). Adjusted durations for the latter are longer probably because the whole word is more sonorant than *sas*, and sonority seems to have a conducive effect on duration perception (an effect we find again in this experiment: vowel duration is more salient than consonant duration). The extremely short adjusted durations, however, remain unexplained. Again we must emphasise that the absolute value of the adjusted durations is of lesser importance to us. We are only interested in adjusted durations relative to that of the base.

The difference between sonorant and non-sonorant codas that we hoped to find is only partially reflected in figure 2. There is no difference between adjustments for *mam* and *sas* codas when we look at shortening, but lengthening definitely has a larger effect for sonorant codas (and onsets) than non-sonorant ones. Also, the fact that the overall adjusted durations for *mam* are longer than those for *sas* points in the direction of a better duration perception for sonorant segments. We conclude that there is marginal evidence for a role of sonority in the perception of consonant durations. Note that, contrary to our expectations, this role can also be played in the perception of onset consonants.

Finally we will have a look at the JNDs, which we again calculate from the standard deviation of the adjustments. In table IV the onset, nucleus and coda JNDs for *mam* and *sas* are given as a percentage of the total duration of the reference stimulus.

Table IV*: JNDs for onsets, nuclei and codas by duration manipulation, for sas (left) and mam (right) calculated from the sd's of the adjustments and expressed as a percentage of the reference duration.*

Duration of Sr	Onset SAS	MAM	Nucleus SAS	MAM	Coda SAS	MAM
270 ms	21	21	18	21	20	20
300 ms	21	19	21	19	21	19
330 ms	19	18	22	20	17	18

As in the pilot, JNDs do not discriminate between onsets, nuclei and codas, nor do they show a difference between sonorant and non-sonorant segments. JNDs centre around 20%, which is high, but not unlike some JNDs for speech reported in the literature (cf. Klatt & Cooper 1975).

In conclusion we can say that the present experiment reproduces the results found in the pilot if we look at it from a broad perspective. The prediction we made with respect to duration perception in onsets was that duration differences in these onsets should be less perceptible than duration differences in either nucleus or coda. This expected difference is found when we shorten the reference syllable. For lengthening we only find the effect for the nucleus to be different from the effect for onset and coda (as a group). Hence, this experiment provides only partial evidence to sustain the hypothesis put forward in section 3.1.

The general effect of subsyllabic constituent type on duration perception is reflected *only* in the fidelity with which the subjects reproduce actual differences of 30 ms in these constituents, *not* in the accuracy with which they perceive duration. The JNDs we find are as large for onsets as they are for nuclei and codas. Remember that we reasoned as follows. When subjects poorly perceive onset duration differences, only major changes in this duration should be audible, which should be reflected in the JND (as opposed to duration perception in nuclei and codas, which should be better than for onsets, and hence, lower the JNDs). In hindsight, however, we fear that the JNDs we are looking for cannot be found through the type of adjustment experiments that we have been running. Contrary to what has been done in the literature, we have, in each case, only altered the duration of *one constituent* of a complex stimulus. It was rather naive to expect to find differences in the JNDs that are a measure for the accuracy with which subjects perceive the duration of the *whole stimulus*. If we could isolate the JNDs for onsets, nuclei and codas, we might find the expected differences. In this experiment, however, the duration changes in the separate constituents could only have a marginal influence on the JND of the total stimulus. The more reliable JNDs from the second experiment (which are based on more subject responses) *all* centre around 20%. Hence, if our duration manipulations modify the overall JND in the predicted direction at all, the effect is probably too small to be detected. But, once again, the JNDs we are talking about here are those for the total stimulus. What we are interested in are the JNDs for the separate constituents. To find these JNDs, and the desired differences between them, we must revert to other experimental paradigms.

Furthermore, the absence of a clear difference between onsets and codas in the lengthening cases is unfortunate. If we wish to draw any conclusions with respect to the phonological status of the syllable onset from a duration perception experiment, we must find a clear effect of a general nature. It might be that the conditions created by the duration of the base syllable and/or the duration variation steps we worked with were not favourable enough to produce the effects we wished to find.

Therefore, a third experiment was run in which more reference stimuli were used to test duration perception for more base durations and duration variation steps in a different experimental paradigm. To improve our chances of finding clear JND differences we chose for the 2I-2AFC experiment that was described in section 3.1.1. In the next section we present it as the final experiment of this chapter.

3.4 Duration perception in a pairwise-comparison task

In the previous sections we have seen that there is some evidence for the claim that the perception of differences in onset duration is imperfect when compared with the perception of such differences in nucleus and coda duration. The main assumption that lies at the base of this thesis is that phonetic duration is the foremost correlate of phonological weight. Any difference in the durational behaviour of certain groups of segments may therefore point at a weight-related phonological difference between these groups. Hence, the provisional conclusion we may draw from the results of the previous experiments is that the relationship between phonological weight and onsets is weak or even absent altogether, while this relationship is stronger for nuclei and codas. We have even found evidence for the role of a secondary phonetic correlate of syllable weight. It appears that the fidelity with which subjects perceive the duration of sonorants is higher than the fidelity we find for equally long continuant obstruents. This result supports the phonological observation that quantity-sensitive stress systems sometimes divide moraic and non-moraic segments on the basis of sonority.

However, the conclusion drawn from the previous experiments is indeed only provisional. The results that were presented in the previous sections were not conclusive. They showed a tendency towards the relative perceptual invariance of onset duration, but in some circumstances onset and coda duration were *both* perceptually invariant. In other cases duration changes in onsets and codas were both perceived with significantly high fidelity (the fidelity for perception of duration variations in nuclei was high, or even tended towards overestimation). Moreover, the poor duration perception for onsets was reflected in the low fidelity with which shortening of onsets was perceived, but not in the accuracy with which duration was perceived. We noted at the end of the previous section that the experimental paradigm in combination with the type of stimuli we used could not yield the constituent-specific JNDs we were looking for. Therefore, the difference we wish to make in the potential weight of onsets and codas cannot yet be based on across the board phonetic evidence.

The adjustment experiments have proven their worth in detecting the asymmetry between duration perception in onsets and codas, and in determining the size of the effect. To find the expected differences in JNDs, however, it would be worthwhile to direct our attention to the 2I-2AFC paradigm. It is our expectation that the JNDs we may find in a 2I-2AFC experiment will reveal the asymmetry in the accuracy of duration perception in onsets, nuclei and codas. Moreover, independent evidence from a completely different experiment strengthens the case for the acceptance of our initial hypothesis.

Finally, in the two previous experiments we have been very limited in our exploitation of the duration factor. The base syllable was always 300 ms, and the duration variations were always 30 ms. We do not know how this affected the results. Therefore, we decided to vary the base duration in the new experiment, at the same time introducing six duration variation steps instead of two. These two new variables may help us to find out whether the duration conditions in the other two experiments were optimal or not. If not, the present experiment will probably reveal the optimal conditions, and what the size of the desired effect is under these conditions.

3.4.1 Method

For this experiment we reused the synthesized word *mam* from the previous experiment. As before, we altered the frame durations of the three segments so that they were all 100 ms.

To generate a set of stimuli we created three base syllables from this LPC-parameter file; one was the original syllable of 3 × 100 ms, while the other two had durations of 3 × 80 and 3 × 120 ms, respectively. The intensity envelopes of these stimuli were adjusted so that the stimuli were almost symmetrical (see figure 3).

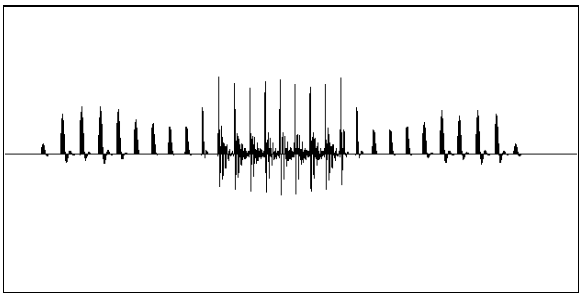


Figure 3: oscillogram of the 300 ms base /mɑm/ stimulus.

The (linear) rise and fall times were set to 30 ms to avoid disturbing clicks. For each of the three base syllables the durations of onset, nucleus and coda were systematically varied through lengthening or shortening in three steps from 20 to 40 to 60 ms. These duration adjustments were carried out through manipulation of the frame duration for the relevant part of the base LPC file.[8] The set of stimuli we thus created is defined by the following possible combinations: base duration {240|300|360} × manipulated segment {Onset,Nucleus,Coda} × duration manipulation {−60|−40|−20|20|40|60}. To the 3 × 3 × 6 = 54 stimuli we added the 3 bases, generating a total of 57 stimuli. The base syllables were used as reference stimuli, while the whole set of bases and derived syllables were used as comparison stimuli. The internal duration make-up of the set of

[8] These adjustments affect the rise and fall times proportionally. Yet, we think this is the better choice from two unavoidable evils. The other option is to delete selected frames, like we did in the pilot. Especially in the extremely short segments this may lead to either inaudible segments if we wish to preserve frames from the rising part of the segment, or even more abrupt rise times if we wish to preserve frames from the steady part of the segment. We thought these effects to be undesirable and opted for proportional shortening. Since all rise and fall times are affected in the same way, the influence of this decision on the relative difference in duration perception between onsets and codas will be negligible.

stimuli generated by base duration{300}, manipulated segment{Nucleus} × duration manipulation{−60|−40|−20|20|40|60} is given in table V as an illustration (compare table I).

***Table V**: durations of individual segments in stimuli derived from the 300 ms base with manipulated nuclei (manipulated segments in bold).*

Duration change	Onset	Nucleus	Coda	Total
−60	100	**40**	100	240
−40	100	**60**	100	260
−20	100	**80**	100	280
base	100	100	100	300
20	100	**120**	100	320
40	100	**140**	100	340
60	100	**160**	100	360

The 57 parameter files thus obtained were converted to sampled data files using LPC synthesis (cf. Vogten 1984).

Sixty-four Dutch subjects participated in the experiment. Their ages varied between 18 and 40, and none of them reported any hearing difficulties. The subjects were paid for their participation. Stimuli were either pairs of a base and a comparison stimulus derived from that base or base-base pairs (to check whether the chance level is 50%), with a silent interval of 500 ms in between. The two possible orderings base-comparison stimulus and comparison stimulus-base were both exploited (to counterbalance the TOE effect, see section 3.2.2), raising the total number of sets to 114. These sets were put in random order and presented to the subjects via earphones attached to an Iris Indigo workstation that was placed in a sound-proofed booth. Three stimulus pairs were presented to the subjects as training material. The subjects were asked to decide which of the two stimuli in a given pair was the longer one by pressing one of two colour-marked keys on the keyboard. It was possible to repeat the trial (by pressing the spacebar). Once a response was given, the computer automatically initiated the next trial. The responses were recorded and stored on disk. The subjects did not receive any feedback on the “correctness” of their response.

3.4.2 Results and discussion

We calculated the percentage of subjects that chose the comparison signal (henceforth C) as the longer one, for each stage of duration manipulation. In the ideal case 0% of the subjects would mark C as the longer stimulus when it is extremely short with respect to the base (or reference stimulus), and 100% would point to C as the longer stimulus when it is extremely long with respect to the base. If C and base are equal, subjects should score at chance level, so that C is chosen as the longer syllable in 50% of the cases. Suppose we visualise the results in a figure with this 50% point, or PSE (point of subjective equality), in the centre of the x-axis, with shortening steps increasing leftwards on the axis, lengthening steps increasing rightwards, and percentages indicated on the y-axis. We then predict that the further we get from this mid-point on the x-axis the less the effect of extra duration steps on the y-axis percentages will be. If duration manipulations are large enough, subjects generally detect the difference easily. Further increases or decreases in duration only lessen the errors, so that judgement percentages come closer to their minimum and maximum limits. Hence, we expect the psychometric functions to run from approximately 0% to 100% in a sigmoid curve (resembling a cumulative normal, or cosine, distribution). Figures below will show three such curves, one for each subsyllabic constituent.

If subjects hear duration manipulations in the nucleus better than such manipulations in either onset or coda, the curves of the latter two should be less steep. Between onset and coda we expect the same difference, onset being less steep than coda. Why this should be so is easily understood. If we lengthen the comparison stimulus, knowing that the manipulation is hardly audible (as in onsets), the percentage of subjects that judge it to be longer than the unaltered base will be lower than the comparable percentage for a stimulus of which the duration change is more salient (as in nuclei). For shortening, the percentage of subjects that judge C to be the longer one will go *down* if the manipulation is more salient. Hence, the nucleus curve should drop lower per shortening step and be raised higher per lengthening step than the coda curve. The same holds true for the coda curve with respect to the onset curve. All this means that the curves should cross each other in the 0 (no-manipulation) point (by definition the curves share the stimuli in this point).

These expectations are borne out by the facts. First we calculated the general percentages, disregarding the differences in base duration and stimulus ordering. These percentages and their respective psychometric curves are presented in figure 4 (exact percentages can be found in appendix B).

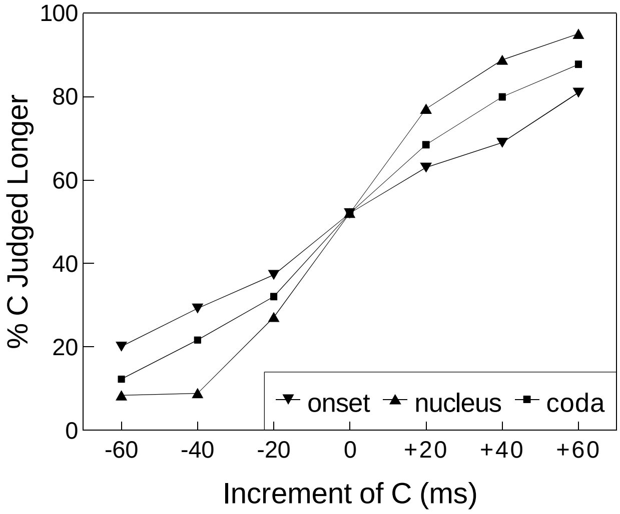


Figure 4: percentages of “longer” judgements for C by duration manipulation and syllabic constituent.

In this figure we clearly see that the nucleus curve is the one that can best be fitted to the prototypical sigmoid function. It starts close to 0% and ends close to 100%. In case shortening is 60 ms most of the subjects correctly judge the comparison stimulus to be shorter than the base. The percentage of “C longer” judgements goes up when C duration comes closer to base duration (so the percentage of correct “C shorter” judgements goes down). At the 0-point (where subjects had to choose the longer stimulus from base-base pairs) we see that responses are at chance level. Both bases are selected in about 50% of the cases, which means this 0-point is indeed the PSE. Past the PSE the percentage of “C longer” judgements in the nucleus curve correctly increases with each duration step, ending near 100%.

In the coda curve the percentage of longer judgements is higher than that for the nucleus when we look at shortening, but it is lower when we look at lengthening. Hence, the number of “errors” is larger for the coda than for the nucleus in all cases. This number of “errors” is the largest in the onset curve, which has the largest percentage of “C longer” judgements in the shortening cases and the smallest number of “C longer” judgements in the lengthening cases. So, the prediction stated at the beginning of this section holds true, degree of correct duration perception is reflected in the slope of the curve, the better the steeper.

As we can see in figure 4 (and in the tables in appendix B), the differentiating effect of duration perception on the three groups is always the biggest in the 40 ms duration change step. This could be taken to mean that 40 ms would have been a better choice for the only manipulation step used in the adjustment experiments. The fact that we have used 30 ms instead may have been the cause of the absence of the desired effect in the lengthening cases of the previous experiment.

The differences between the curves for onset, nucleus and coda prove statistically significant. For each of the curves a slope-coefficient was calculated through fitting the sigmoids to a regression function. These coefficients were then tested for equality in a t-test. Comparing the values for the nucleus and the coda curve we find $t(6) = 6.4$, $p < .005$, and for the coda × onset we find $t(6) = 5.2$, $p < .005$.[9] We may conclude, therefore, that this experiment confirms the effects we found in the pilot experiment. The perception of duration is different for onsets, nuclei and codas. Again we find a significant difference between onsets and codas. The difference between codas and nuclei, which we provisionally attribute to sonority, is also confirmed.

A characteristic of this type of experiment is that the JNDs for duration perception can be directly deduced from the psychometric curves. Thus we can determine the JNDs for onset, nucleus and coda in an easy way. By agreement, the JNDs are defined by the abscissa corresponding to the point in the curve at which 75% of the subjects correctly determine the difference between the reference and the comparison stimulus (cf. section 3.1.1). In our case we must look at two points in each curve. For lengthening we must look at the abscissa corresponding to the 75% point. At the 25% point, however, 75% of the subjects correctly judge C to be shorter, which gives us the JND for shortening. In table VI these JNDs are presented together with the mean JND collapsed over lengthening and shortening.

We conclude that this 2I-2AFC experiment has yielded the difference in JNDs we had hoped to find. Poor duration perception in onsets is reflected in the relative perceptual invariance of its duration, as evidenced by the previous experiments, but also by the accuracy with which duration variation is detected. The JNDs presented in table VI mirror the effects found in the pilot experiment. They are large for onsets (poor detection), smaller for codas and smallest for nuclei, just as we predicted. Note also that the percentages for vowels and coda consonants closely

[9] Statistic sources differ on the degrees of freedom that apply to this type of test. We choose the most stringent option and consider this to be a t-test for paired samples. The degrees of freedom are then: 7-1=6 (instead of 12 in the other type). We have found significance in the stricter test, so we would also have found it in the less strict version. The t-test is one-tailed since we predict the slope of the onset to be less steep than that of the coda, while the slope of the coda should be less steep than that of the nucleus.

resemble those found for Japanese synthetic speech (vowels 7%, /s/ 10%) by Fujisaki et al. (1975).

***Table VI**: JNDs for onset, nucleus and coda in the lengthening and shortening case in ms. Also presented is the mean JND (collapsed over lengthening and shortening) expressed as a percentage of the mean duration of the reference syllable.*

Duration change	Manipulated Constituent Onset	Nucleus	Coda	Reference Duration
shortening	48.9 ms	22.2 ms	33.3 ms	300 ms
lengthening	48.1 ms	17.4 ms	30.4 ms	300 ms
mean	16.2 %	6.6 %	10.6 %	300 ms

These findings do not conclude the list of results we may obtain from this experiment. Remember that figure 4 represents the grand mean that was calculated while ignoring two independent variables that we put in the experiment: we varied the base duration and the position of C in the stimulus pair. Let us first look at the base duration. We calculated the “C longer” percentages for each base separately and plotted the curves (the resulting figures can be found in appendix B); we then determined the JNDs at the 25% and 75% points for the three subsyllabic constituents in these bases. The mean JNDs (collapsed over lengthening and shortening) for the three different bases are given in table VII.

***Table VII**: mean JNDs for onset, nucleus and coda in the three base duration cases (collapsed over lengthening and shortening) expressed as a percentage of the duration of the reference syllable.*

Base duration	Manipulated Constituent Onset	Nucleus	Coda
240 ms	18.6	8.8	9.8
300 ms	15.7	6.2	10.1
360 ms	13.6	5.8	10.1

These JNDs resemble those found in table VI. There are some minor differences between tables VI and VII which we attribute to the fact that each of the JNDs in table VII represents only a third of the total data pool we used to calculate the JNDs in table VI. For each base we find the expected difference between the JNDs of onset, nucleus and coda.

Notice that the JND for the coda remains very close to the 10% that is generally reported in the literature. The other JNDs in table VII go down as the reference duration goes up. Apparently subjects perceive duration variations in longer onsets and nuclei better than such variations in shorter segments. The conditions for accurate duration perception, though, seem to be most favourable with the 300 ms base. JNDs are generally lower than for the 240 ms base, and the difference between the percentages for minimal and maximal JND is higher than for the 360 ms base (the difference between onsets, nuclei and codas is maximal). It seems we were right in choosing the 300 ms base in the previous experiments, since this base benefits the accuracy of duration perception.

The final results are obtained through the position factor. The two orderings of the base and C were only put in to counterbalance the effect of the time-order error (TOE: see section 3.2.2) by which the duration of a final stimulus is underestimated (or the duration of the prefinal stimulus is overestimated). However, by looking at the PSEs of the two orderings individually we can gauge the size of the TOE in our experiment and present it to complete the list of results for this experiment. When the TOE is 0, subjects select each stimulus of a pair of equally long stimuli as the longer one in 50% of the cases: in this trivial case the point where the stimuli sound equally long (PSE) is reached when reference and comparison stimulus are indeed equally long. For the base-C ordering in our experiment, however, the subjects score at chance level (50%) when C is 4 ms *longer* than the base (see figure 6 in appendix B). That means the duration of C is underestimated. For the C-base ordering, the subjects score at chance level when C is 7 ms *shorter* than the base (see figure 6 in appendix B). In this case the duration of the base is underestimated. We conclude that underestimation of the final (or overestimation of the prefinal) stimulus, according to the TOE, is present in our experiment. The TOE effect, found by the subtraction of the values on the x-axis corresponding to the PSEs for both orders, is 11 ms. It appears we have done well to eliminate it by varying the order of the base and the comparison stimulus. This also confirms the suspicions stated in section 3.2.2 (just below table II). The extremely short adjustments we found in the adjustment experiments were not due to the TOE. For one they were much too large, and secondly, they pointed at an *over*estimation of the final stimulus. These findings run counter to the real TOE effects that are uncovered here, which replicate those reported in the literature. The overestimation effects found in the adjustment experiments remain elusive, but are most likely caused by the differences in reference and comparison signal (speech vs. non-speech).

3.5 Conclusion

The experimental results presented in this chapter are quite different from those we found in the previous chapter for duration in speech production. This time we have concentrated on the perception of duration, and we do find a strong asymmetry between onsets and codas. The first two experiments revealed that the absolute perceived duration of onsets remains relatively invariant with respect to that in codas. Differences in onset duration are perceived as less sizeable than equal differences in coda duration. These experiments also showed that the differences in nucleus duration were even more salient than the differences in coda duration. We expected to find the same division in the JNDs for duration. If duration changes in onsets are perceptually underestimated, the minimal duration change that can be perceived in onsets will most likely be more sizeable than the minimal duration change that is still perceptible in codas (or nuclei). Hence, we expected onset JNDs to be higher than coda JNDs which, in their turn, should be higher than nucleus JNDs. We did not find this difference in the first two experiments, but attributed this to our experimental design. The third experiment was conducted to confirm the relative invariance of onset duration found in the first two experiments, and to check whether the poor duration perception in onsets is also reflected in their JNDs. In this experiment, in which we used a different experimental method, the results also indicated relative perceptual invariance of onset duration. Moreover, this time we did find the expected differences in JNDs.

3.5.1 Linguistic consequences

If we take durational asymmetries as evidence for differences in potential phonological weight, we must now explain a two-way weight opposition. The difference between onsets and codas seems to confirm the phonological view that codas may be dominated by a moraic node while onsets may not. By the same view we know that nuclei may also have a mora, which is reflected in the excellent perception of duration in vowels with respect to both onsets and codas. The fact that perception in nuclei is better than perception in codas is not a problem. Metrical systems in many languages consider only the length of the vowel when they determine syllable weight. Languages in which the coda also adds to weight employ an extra rule (WBP cf. section 1.2.2) to invoke the potential weight of the coda. The difference in the perception of duration in nuclei and codas that we have uncovered may reflect the difference between default weight units (vowels) and add-on weight units (coda consonants). Alternatively, we may look at sonority to explain the difference. Vowels and consonants obviously have different sonority values. This sonority factor often plays a role in stress rules (see Hayes 1995; Goedemans 1996a). The general tendency is that segments with a higher sonority are more likely to add weight to a syllable. The extra difference we find in the duration perception of the two moraic segments (nucleus and coda) may be caused by the added contribution of sonority. Not only is duration perception better in segments that may have a mora, it is also better in sonorant segments than in non-sonorant segments. Note that the latter does not exclude the onset. We have also found duration perception to be better in sonorant onsets. This possible sonority influence on duration perception in onsets, together with the fact that duration perception in onsets is only relatively invariant (differences in onset duration can be perceived, though poorly), force us to conclude that a contribution of onsets to syllable weight is not impossible. Languages may exist in which onsets can make a syllable heavy, or heavier, though the poor contribution the onset can make with respect to the contributions of codas and nuclei makes this extremely unlikely (see chapter 5 for further discussion).

In any case, we may safely assume that perceived duration is the primary phonetic correlate of phonological weight. The weightlessness of the onset is reflected in its relative durational invariance and its high JND. However, these findings do not allow us to accept the hypothesis stated in section 3.1 completely. The hypothesis stated that the relative perceptual invariance of onset duration *explains* its weightlessness. Yet, we do not know whether the effects we have found are of such a general nature that they force upon our internal representation of the syllable the restriction that onsets must be weightless, or whether a representation of syllables without onset weight causes our imperfect perception of onset duration. In the next chapter we will try to find an answer to these questions.

5 The Role of Onsets in Stress Rules[1]

5.1 Introduction

With this chapter we start the second part of this study. The change is not so much in the topic of our research as in the field of linguistics that is highlighted. In the previous chapters we have been mainly concerned with phonetics, while in this, and the next chapter the focus will be on the phonology of onsets and their role in stress rules. The necessity for this phonological excursion became clear in an early stage of what would have been a largely phonetic project. It came to our attention that the universality of the rule that onsets are weightless was disputed by, among others, Davis (1985). In his dissertation, he mentions several languages in which onsets influence the assignment of stress. According to him, these languages form counterexamples to the claim that onsets cannot contribute to syllable weight. As such, he uses these languages in his argumentation for a flat syllable structure in which onset, nucleus and coda are all potentially weight bearing units which are directly linked to the syllable node.

From our point of view, the existence of these would-be moraic onsets casts doubts on the validity of the research question we started out with (i.e. *Why is the syllable onset weightless?*). If onsets are not always weightless, then we should at least conclude that there is no absolute phonetic principle that causes such weightlessness. Early results from the pilot experiments described in chapters 2 and 3, however, showed evidence for a systematic weight related difference in the perception of onsets on the one hand, and nuclei and codas on the other. Supported by these results we continued our experiments, at the same time starting a search for phonological alternatives to the moraic-onset analyses that were presented by Davis (1985) for the languages in question. In view of the experimental results that we presented in the previous chapters, in which we confirm the weightlessness of the syllable onset, the possible

[1] Large parts of this chapter are adapted from Goedemans (1996a).

existence of moraic onsets becomes all the more dramatic, and the search for alternatives all the more urgent.[2]

During this search for alternative metrical analyses it appeared that, besides the ones Davis mentions, more languages exist in which onsets seem to play a role in stress assignment. In this chapter we give an overview of these languages, dividing them into two groups. One group contains all the languages for which a straightforward moraic onset seems to be needed. For these languages some historical background is presented, and the way is paved for a non-moraic reanalysis of their unusual stress systems. In section 6.2 such an analysis is presented for two related languages from this group: Western Aranda (Strehlow 1942) and Alyawarra (Yallop 1977). It is our belief that this analysis applies to most (if not all) of the languages in that group.

The second group contains languages that seem to exploit the possibility of onset prominence that we touched upon at the end of the previous chapter. Remember that we did not exclude the existence of stress systems referring to onset prominence. Contrary to an abstract mora, prominence factors may be present in the onset of the prosodic syllable. Though it seems unlikely, languages may use these factors in a heavy/light distinction. However, we also noted that there are restrictions on the type of factor that may play a role in prominence relations. Some of the languages we will discuss employ true onset prominence in their stress systems, while the influence of onsets on stress assignment in some other languages cannot be attributed to prominence. Even the proponents of moraic onsets cannot use moras in an analysis of the latter type of languages, because the (apparent) distinctions between heavy and light syllables in these languages do not involve the presence or absence of an onset, but of a feature in that onset (by the same token, they cannot use moras in an analysis of a stress system that uses coda sonority, but see section 5.2.3). Hence, they are problematic in any metrical theory. However, the possibility exists that these languages do not show any influence of onsets on stress rules at all, and that the observed phonological regularities concerning onsets and stress are coincidental. In-depth phonological studies of these languages may reveal the hidden sources that cause the stress rule to be mistakenly labelled as "onset-sensitive". This seems to be the case for Mathimathi (Hercus 1986), which is the subject of a case study we present at the end of chapter 6.

Before we start our overview of languages with onset-sensitive stress

[2] Albeit, at the end of chapter 4 we stumbled upon some evidence for the claim that abstract linguistic principles play a role in the assignment of moras, which means that one cannot exclude a linguistic rule that assigns a mora to onsets. However, we consider the existence of such a rule to be highly unlikely.

rules, some background information on metrical phonology will be provided in the next section. The examples contained in this section are drawn from Australian Aboriginal languages. Since the majority of the onset-sensitive stress systems that we encountered can be found on the Australian continent, we thought it wise to gather as much information on stress systems in Aboriginal languages as we could find, so as to sketch the “metrical setting” for the systems that are relevant to us. This “setting” is provided as an annotated list of languages in appendix D.[3] Through reference to that list (while exhibiting examples from it here and there) we hope to improve the internal cohesion of this and the next chapter. This procedure should also serve to convey a sense of familiarity with Australian metrical data, such that the discussions below and in chapter 6 can be placed in their proper context.

5.2 Metrical phonology

As has been noted in chapter 1, metrical phonology is the field of linguistics concerned with the rules that are needed to derive the *fixed* stress patterns we find in natural languages (aided by the rhythmic properties we also find in music and verse). In this section we will briefly sketch the main principles of metrical theory. We will refer to the metrical rules in a very general way. No definitive choice will be made for a particular formalism, though the examples will be presented in a fashion that closely resembles the framework proposed by Hayes (1995). Only in chapter 6 the choice for a specific metrical formalism will be made. The basic metrical rules, and some common stress patterns, will be introduced on the basis of stress data from Australian languages. Each example will be accompanied by a discussion of the rules needed to derive the observed stress pattern, or the relevance of the example for metrical theory in general. In this way we will cover all that is needed to obtain a basic working knowledge of metrical theory. The intricate details that are relevant to the languages that have onset-sensitive stress rules will be introduced during their respective discussions.

---

[3] Most of this work was done in Australia at the Australian National University and the Australian Institute for Aboriginal and Torres Strait Islander Studies. Both institutes are gratefully acknowledged for their hospitality. The references to the descriptive sources for the languages given in appendix D are not repeated in the reference section at the end of this book.

5.2.1 Quantity-insensitive stress

The approach that metrical theory takes with respect to the analysis of stress patterns is the following: for every language a structure of units consisting of weak and strong nodes is derived that can be placed over the words, predicting the locations of the stressed (strong) and the unstressed (weak) syllables. The structure is not arbitrary; it must be derived through the usage of a limited set of well defined binary parameters that define the shape of the structure by their settings. In this section we will introduce the basic parameters: *Foot-type, Boundedness, Iterativity, End Rule, Quantity-sensitivity, Directionality* and *Degenerate feet*, in that order.

The prototypical stress pattern for an Australian Aboriginal language is to have main stress on the first syllable and secondary stresses on every odd syllable thereafter (sometimes excluding the final syllable). The Gugada dialect of Western Desert (Platt 1972) is an example of such a language. In (1) we present some examples (*á* denotes main stress, *à* denotes secondary stress).

(1)	bádu	'man'
	wáljabàra	'whitefellow'
	ánguŋàrinjdjàgu	'want to sleep'

From a theoretical point of view the Gugada stress pattern is unproblematic. In fact it represents the unmarked option in most theories. First of all we observe that stresses occur at regular intervals. Judging from the pattern, it seems logical to build a structure that divides the word into several binary rhythm units that all contain two syllables. The units that are handed down to us from verse are *feet*. The *Foot-type* parameter allows two basic varieties: *iambs* and *trochees*. These feet represent the relative strength of the two syllables contained in them, labelling one of them as weak, and the other as strong. The strong syllable is usually called the head of the foot. It is not difficult to guess the variety in their internal composition: trochees are left-headed and iambs are right-headed (represented as (* .) and (. *), respectively). Since the first word in (1) exactly reflects the strong-weak pattern, we may assume that we need trochees to derive the Gugada stress pattern.

The Gugada pattern is characteristic of a *bounded* language; the main stress is located at one of the word edges, and the feet needed to parse the words are binary. In *unbounded* languages, stresses may occur anywhere in the word (yet they are predictable by rule, see section 5.2.2) and there is no limit to the size of the feet.

A logical type of structure that we can build with the parameters we have introduced so far is one that places one iamb or trochee at one of the word edges and then stops, deriving one (main) stress only. These structures are useful indeed, because stress rules of such a limited nature do exist in natural languages. The parameter that regulates this persistent or non-persistent assignment of feet is called *Iterativity*. It is set to “no” if only one foot has to be built. Since more than one stress is present in two of the words in (1) we need to place feet throughout the entire word to derive the Gugada patterns. Hence, we label foot assignment in Gugada as iterative.

Finally, we observe that one of the stresses is always branded as the strongest one; the main stress. We obviously need a parameter that allows the structure to differentiate between main and secondary stress. Remember that we predicted the occurrence of different stress levels in chapter 1 (footnote 7 in section 1.2.1). Only one syllable can bear the accent when the word is in focus, though iterative footing of words that are large enough marks several syllables as strong. To select from those strong syllables the one that will carry main stress we use the *End Rule* (right/left) to promote either the rightmost or the leftmost foot-head to a second (main stress) level in the structure. In the Gugada case we need End Rule (left). Combining the settings of the parameters discussed above we can predict the stress patterns of all Gugada words. In (2) the final word from (1) is repeated with its metrical structure and the parameter settings needed to derive that structure.

(2)	(*)	main stress or word level	Bounded: yes
	(* .)(* .) (* .)	foot level	Foot type: trochaic
	ángu ŋàrinjdjàgu		Iterative: yes
			End rule: left

The iteratively assigned trochees cover the word completely. Each foot-head dominates a stressed syllable and the first foot-head is selected as the most prominent one and represented on the word level. It appears that the trochees in (2) are concatenated rigorously, one for every two syllables, irrespective of the internal make-up of these syllables. Remember from the discussion in section 1.2.1 the term *quantity-insensitive* (QI) that we introduced for languages like Gugada, in which the internal composition of the syllables is irrelevant. Languages can also be *quantity-sensitive* (QS), but we will refrain from reference to this parameter here, since we discuss QI and QS languages in separate sections (this section and 5.2.2, respectively), keeping the settings of the parameter constant in each section.

A slightly more complicated stress rule is exemplified by Arabana-Wangkangurru (Hercus 1994). This language has the same basic pattern as Gugada, but when words are morphologically complex the pattern is repeated for each non-monosyllabic morpheme:

(3)	wárpa	'storm'
	wánparda	'carry'
	yúrkuràngku	'the ancestral black snake'
	kátha-nàngku-lìparna	'they used to travel continually long ago'
	kátha-rnda-nàngku-lìparna	'they used to travel continually and in a hurry long ago'

In many other languages such a repetetive stress pattern is also reported to occur on *suffixes* that are disyllabic or longer. If we wish to capture these facts in a metrical theory we need to allow for rules that refer to these morphological categories, a feature that all theories have. (In appendix D the label (M) is added to languages in which morphology plays a role.)

With the help of the word *wánparda* we can illustrate another important parameter in the analysis of stress patterns. Note that we need trochees to derive the Arabana stress pattern, and that we cannot just choose at which word edge we start building the feet. In *wánparda* we must start at the left edge, otherwise we derive **wanpárda*. The *Directionality* (left/right) parameter is the one that determines at which edge footing must start. The parameter settings for Arabana are the same as those given for Gugada in (2), expanded with Directionality (left). When we build the structure for *kátha-nàngku-lìparna* we get (4).

```
(4)   ( *                   )
      ( *   .)(*     .)(* . )
      kátha-nàngku-lìparna
```

Note that the feet built over the three morphemes are integrated into one word-level constituent.

A more important observation concerns the final morpheme, in which the final syllable is not footed. This is obviously due to the fact that no disyllabic foot can be built when only one syllable remains. Yet, many languages do stress such final single syllables, which means that they must be incorporated into the foot structure in those languages. Consider the data from Murrinh-pata (Walsh 1976) that are given in (5).

(5) wéɹe 'dog'
nígunú 'she'
kánaŋándaŋ 'emu'

We observe that we need to place a foot over the final syllable of the second word because it carries stress. This final foot, though, is necessarily incomplete since there is no syllable to the right of the head that can occupy the weak position of the trochee. We need these so-called *degenerate feet* for all languages that share the Murrinh-pata pattern. Languages like Arabana do not stress the final syllable. This hints at a parameterised prohibition/allowance of degenerate feet.[4] For Murinh-patha the parameter is set to allow degenerate feet, which leads to the representation in (6).

(6) (* .)(*)
nígu nú

The data from Murrinh-patha are also important in another respect. Walsh notes there is no difference between the strength of the stresses in Murrinh-patha words. Although we indicate the stresses in (5) with the accent we normally use for main stress, we follow the mainstream of metrical theory in the claim that there is no main stress in languages like Murrinh-patha, which is not unique in this respect. Such languages simply lack an End Rule that promotes one of the stresses to the word level (indicated by NMS, "no main stress", in appendix D).

Languages of the types discussed above are very common in Australia. They occur all over the continent and seem to represent the unmarked stress pattern for Aboriginal languages (they are listed under I and II in appendix D).

Let us now explore somewhat further the possibilities that are opened up by the parameters we have discussed so far. Suppose we were to change the directionality parameter from left to right, and leave all other parameters unchanged. That would give us a system which builds trochees from right-to-left and which promotes the head of the last foot it assigns to main stress. Since feet cover two syllables at a time, the remarkable stress pattern we thus derive stresses the first syllable of words with an even number of syllables and the second syllable of words with an odd number of syllables (provided degenerate feet are forbidden). Such systems do indeed exist, but they are rare. After van der Hulst

[4] For a discussion on the theoretical importance of this parameter and its relation to the occurrence of monosyllabic words, see Hayes (1995).

(1996) we call these systems *Count Systems* since they seem to “count” the number of syllables before assigning main stress. An Australian example is Ngankikurrunggurr (Hodinott & Kofod 1988). The pattern and analysis are demonstrated by the two words in (7) (cf. III in appendix D).

```
(7)   (*        )          ( *           )         Bounded: yes
      (* .) (*  .)         ( *   .) ( * .)         Foot type: trochaic
       éfe kìmi            anímpirrmìre            Direction: right
      ‘rabbit’             ‘firefly’               Iterative: yes
                                                   End rule: left
                                                   Deg. feet: no
```

The opposite kind of count system (left-to-right foot assignment and End Rule right) cannot be found among the Aboriginal languages. The rarity of count systems constitutes evidence for a claim concerning the two parameters that are involved here: Universally there is a strong correlation between the edge at which footing starts and the edge at which main stress is located. So, to derive some Aboriginal stress patterns that are more frequent than count systems, we might set the End Rule parameter in (7) to the same value as the Directionality parameter. Thus, we derive penultimate stress patterns, as is shown in (8) for some Ngalakan (Merlan 1983) words. Trochees are built from right-to-left and the rightmost one is to carry the main stress.

```
(8)   (    * )                          (      *    )
         ( * .)                         (*  .)( *  . )
      burkáji     ‘genuine, real’       mìli bálkiñ   ‘salt water’

      Bounded: yes                      Iterative: yes
      Foot type: trochaic               End Rule: right
      Direction: right                  Deg. feet: no
```

Ngalakan is spoken in Arnhem Land (Northern Territory). Among the languages spoken there, quite a few do not obey the general tendency for Aboriginal languages to have some kind of left oriented stress rule (see IV, appendix D). Instead they place main stress on the penultimate syllable. Where this deviant stress pattern historically comes from is unclear. It may represent an old pattern that was more widely spread in the past but has been replaced by the initial patterns in languages of more dominant tribes. It may also have arisen only recently through contact with other languages. The predominantly penultimate stress patterns of the languages in the nearby Indonesian area are clearly suggestive. In any case, languages with main stress on the penultimate syllable form a minority group among Aboriginal languages. Unlike count systems, though, they are frequently attested world-wide.

One parameter we have not fully exploited yet is Iterativity. Normally, non-iterative stress assignment is uneventful. One foot is assigned at the left or the right word edge, the head of which is obligatorily promoted to main stress. More interesting patterns, in which non-iterative foot assignment is combined with normal iterative foot assignment, can be found in Garawa (Furby 1974) and Nunggubuyu (Hore 1981). Those patterns are illustrated in (9).

(9)	*Garawa*	
	a. púnjala	‘white’
	wátjimpàŋu	‘armpit’
	yákalàkalàmpa	‘loose’
	b. kámalařìnji	‘wrist’
	ŋánkiřikìřimpàji	‘fought with boomerangs’
	Nunggubuyu	
	c. wurúgu	‘billabong’
	ngàlaalígi	‘turtle’
	màragàrrijínyung	shark species
	d. àmbalalári	‘poor’
	ràwurrùmugurrúmu	plant species

Judging from the data in (9a) and (9c) Garawa exhibits the Arabana (initial stress) pattern and Nunggubuyu looks like Ngalakan (penultimate stress). For main stress the resemblance holds true across the board. However, the Nunggubuyu words veil the fact that a secondary stress occurs on the initial syllable even if iterative footing from right-to-left would result in a secondary stress on the second syllable, as is shown in (9d). The Garawa words in (9b) show that there is always a penultimate secondary stress, irrespective of whether the number of syllables between it and the main stress is odd or even. These uncommon patterns can be derived if we recognise Garawa and Nunggubuyu as *bidirectional* systems. What happens is that a non-iterative foot is assigned at one of the edges while iterative footing starts at the other edge. The non-iterative foot is the one that is promoted to main stress. So, the iterativity and directionality parameters must be set twice, once for the main stress foot and once for the secondary stress feet. In (10) we give the parameters for both languages and provide the structures for two relevant words (both languages are bounded and trochaic, and forbid degenerate feet).

(10) *Garawa* | *Nunggubuyu*

Garawa	*Nunggubuyu*
Direction: left + right	Direction: left + right
Iterative: left no, right yes.	Iterative: left yes, right no.
End Rule: left	End Rule: right
(*)	(*)
(* .) (* .)(* .)	(* .)(* .) (* .)
ŋánki řikìřimpàji	ràwurrùmugurrúmu

Note that the syllable that is left over after iterative footing cannot be parsed into a foot because degenerate feet are disallowed (as in Arabana and in the first words in (9a) and (9c)).

Languages like Garawa and Nunggubuyu are of considerable theoretical importance. The fact that they seem to need two modes of parsing is used by van der Hulst (1984) as evidence for a theory in which the assignment of main and secondary stress are separated. Furthermore, these languages are claimed to be intermediate between left-to-right systems and right-to-left systems (as are possibly also count systems like Ngankikurrunggurr and Malakmalak, Birk 1976). The fact that these four languages are all spoken in Arnhem Land, where languages with the regular penultimate type of stress border languages with the Gugada (initial) type of stress, is clearly suggestive of a borrowing situation.

Apart from the boundedness parameter, which we will discuss in the next section, the only parameter we have not yet switched in our discussion is Foot-type. Stress patterns for which we could use iambs in the metrical analysis can be found in four languages on the east coast of Australia (see V in appendix D). In (11) some examples from Gureng-Gureng (Holmer 1983) are presented that exhibit the pattern these four languages share.

(11) gunáŋal ‘frightened’ gilámanmin ‘turned around’

A possible analysis places one non-iterative iamb over the first two syllables. This analysis would be suspect for several reasons. First of all, in contemporary metrical theory the existence of QI iambs is denied (see, among others, Hayes 1995). Some even go as far as abolishing all iambs completely (van de Vijver 1997). Secondly, in view of the evidence presented in appendix D, we can state that iambic stress is probably not a feature of Aboriginal languages.[5] As holds true for the languages in the

[5] Except possibly for Yidin^y (Dixon 1977). But even there the iambic pattern seems to be enforced by other factors. If these factors are absent, the basic pattern is trochaic.

rest of the world, QI iambic patterns can be reanalysed trochaically. We only need the help of a device called *extrametricality* (EM), the existence of which is independently motivated by its indispensable role in the analyses of other metrical patterns (see Liberman & Prince 1977; Hayes 1979). By EM a certain prosodic constituent at one of the word edges is made invisible to the metrical rules. Suppose we apply EM to the first syllable of Gureng-Gureng words and then build a non-iterative trochee. Thus, we derive the structure in (12) (<> denotes EM).

```
(12)  (      *          )
          (*   .)
      <gi>lámanmin
```

In this fashion all QI iambic systems can be reanalysed, and we can save the hypothesis that all Australian Aboriginal languages are trochaic.

So far, we have only looked at QI stress rules. All the parameters that were introduced, however, can also be applied to QS languages, in which the elements in the rhyme part of the syllable influence stress placement. We will discuss some of these languages in the next section.

5.2.2 Quantity-sensitive stress

Stress rules that refer to syllable weight are far less common among Aboriginal languages than rules that ignore this factor. In some languages we find only marginal references to weight. Consider, for instance, Gaanay, for which Hercus (1986) gives the following rule for secondary stress assignment (main stress is on the first syllable): a secondary stress occurs on the final syllable if it is closed. In (13) we present two pairs of words that illustrate the crucial difference.

(13)	ŋáraŋda	‘to bury’	bíndjulàŋ	‘cat’
	náŋera	‘saltwater mussel’	jálamàn	‘several’

Derivation of the Gaanay pattern requires a small adaption of the Gugada rules. This time we use QS trochees. In contemporary metrical theory, QS trochees are feet that respect the weight of syllables by being disyllabic when placed over two light syllables, but monosyllabic if placed over a heavy syllable. Single light syllables cannot be parsed if we use only QS trochees. This is demonstrated in (14).

```
(14)  ( *     )     (*      )
      ( * .)        (* .)( *)
      náŋera        jálamàn
```

Note that the last foot assigned to *jálamàn* is not degenerate. QS trochees that dominate a heavy syllable are full-blooded feet, which becomes clear if we realise that they must not dominate two syllables, but two moras. Degenerate feet are prohibited in Gaanay, as evidenced by the analysis of *náŋera*.

A problem with this approach is formed by the fact that it differentiates between the assignment of main and secondary stress as far as quantity-sensitivity is concerned. In (13) we see that Gaanay, among others, assigns main stress in a QI fashion while secondary stress is QS. If we do respect quantity when assigning main stress we derive main stress on the second syllable of *ŋáraŋda* since that one is heavy (the first light one is skipped because degenerate feet are forbidden). Clearly main stress assignment must be QI, since Gaanay main stress is always on the initial syllable. Such a QS/QI difference between main and secondary stress, as well as the differences in directionality and iterativity of main and secondary stress that we noted in the discussion on Garawa and Nunggubuyu, point in the direction of a separation between the algorithms we use to assign those stresses. This line of reasoning is fully exploited by van der Hulst (1984, 1996a, 1996b).

Theoretically less demanding stress systems are those that use QS feet for both main and secondary stress. One of the languages that has such a system is Gumbaynggirr (Eades 1979), though reference to QS feet for secondary stress is void in this case because heavy syllables never occur in positions where they would receive such a stress; they only appear in either the first or the second position of the word.[6] Main stress in Gumbaynggirr is on the first syllable unless the second syllable is heavy. Heavy syllables are those with a long vowel or a vowel - semi-vowel sequence. Some examples are given in (15).

(15)	ŋámi	'woman'	ŋalíː	'1st person DU INC'
	míːmi	'mother'	gamáy	'spear'
	ŋáliwan	'jewfish'	ŋalúːŋgir	'clever man'

As noted above, QS trochees are monosyllabic when they dominate a heavy syllable, and they are disyllabic when the head dominates a light

[6] This is a characteristic that is shared by most Aboriginal languages that have a stress pattern similar to that of Gumbaynggirr (such languages can be found in appendix D, VI).

syllable. So, heavy syllables cannot be dominated by the weak part of the foot. Judging from the word *ŋáliwan*, in which we find no final secondary stress, degenerate feet are forbidden and foot assignment is left-to-right. From these data we cannot infer settings for boundedness or iterativity. Below, some of the examples are repeated with their metrical structure.

```
(16)  (*      )    (  *    )   (*   )       (  *)       (*    )
      (* .)          (*)       (*  .)         (*)       (*)
      ŋáliwan      ŋalúːŋgir   ŋámi         ŋalíː       míːmi
```

Other languages have stress patterns that are similar to that of Gumbaynggirr, but these may have long vowels in other positions besides the first two syllables. Metrical rules may exploit this feature and allow main stress to occur anywhere in the word. The rule, as exemplified by some Yukulta examples in (17) (taken from Keen 1983), is just an *unbounded* variant of the rule discussed above. Main stress in Yukulta falls on the first long vowel (heavy syllable) and, if there are no heavy syllables, on the first vowel. Secondary stress is on the penult if possible (i.e. a disyllabic foot can be built that is not adjacent to the main stress).

(17)	ŋíta	'fire, firewood'	ŋúːrpaḷùt̪a	'to finish'
	mákuwa	'woman'	pit̪íːnta	'boy toddler'
	kúlurùna	'bushfire'	puŋkalaṇṭíːtya	'to kneel'

In this case the domain for main stress assignment is the whole word. An unbounded foot (possibly spanning more than two syllables) is built, beginning at the left word edge. Primary stress is assigned to the first heavy syllable in that foot or to the leftmost one if there are no heavy syllables. Notice, however, that secondary stress seems to be QI and bounded, it always occurs on the penult if possible and "eats" two syllables off the unbounded foot. Hence main stress is QS and unbounded and secondary stress is QI and bounded. Again we find a language that supports a separation of the main stress and the secondary stress algorithms. Some of the words in (17) are repeated in (18) with their metrical analysis.[7]

---

[7] In the structure for *mákuwa* there is room for a binary secondary stress foot, as in (i):

```
(i)    (*     )                      (*     )
       (*)(*  .)    destressing      (*)(   .)
       mákuwa          →             mákuwa
```

The secondary stress is deleted because it *clashes* with the main stress. Both this analysis and the one in (18) are possible, and we do not choose between them. Note that this option is not open for *pití:na* since there is no room for a second foot there.

```
(18)  (*      )   ( *        )  (*           )      (  *    )
      (*  .   .)  ( * .)( *  .) (*     .)( * .)     ( . *   .)
      mákuwa      kúlu rùna     ŋúːrpa ḷùṯa         piṯíːnta
```

Bounded: no + yes	Iterative: no
Foot type: QS + QI trochaic	End Rule: left
Direction: left + right	Deg. feet: no

Note that Yukulta is a very special type of unbounded system. Usually we find systems that place main stress on the first or last heavy syllable and secondary stress on all other heavy syllables. In these cases both main and secondary stress are unbounded and QS. Systems like this can be found in the southeast of Australia (see VII in appendix D).

The final type of stress rule we discuss is reminiscent of the rule we found in Ngankikurrunggurr in the previous section. In Warrgamay and Nyawaygi, spoken in the Cairns rainforest region, the typical stress pattern for right-to-left count systems is complicated by the fact that long vowels may appear in initial syllables (only) and that the stress rule seems to be sensitive to them. Let us look at some Warrgamay examples (from Dixon 1981).

(19)	múːba	‘stone fish’
	gagára	‘dilly bag’
	gígawùlu	‘freshwater jewfish’
	guɽágay-mìri	‘Niagara Vale-FROM’

These examples can all be handled in the same way as the examples in (7). The deviation from that pattern appears in words with an odd number of syllables that have a long vowel in the initial syllable, like *gíːbara* ‘fig tree’. These words get initial main stress, though rigorous application of the Count System stress rule would derive main stress on the second syllable. If we assign QS trochees, however, the first (and only the first) syllable may receive a monosyllabic trochee if it contains a long vowel. Words like *gíːbara* would receive a metrical structure like that in (20).

```
(20)  (*     )         (*     )
      (*)(* .)         (*)(   )
      gíːbara    →     gíːbara
```

The secondary stress on the second syllable is deleted because it clashes with the main stress.[8]

This concludes the discussion on the regular stress patterns we may encounter among Aboriginal languages. The parameters we have introduced allow for many more possibilities in the stressing of words, many of which are exploited by the languages of the world (though not in Aboriginal languages). All these patterns can be derived by application of the parameters we have introduced.[9]

Remember that, at the end of chapter 4, we opened the possibility of stress rules that were sensitive to other syllabic properties than weight alone. In the next section we will discuss some languages for which we seem to need such rules.

5.2.3 Prominence systems

As noted above, some languages do not refer to the structure of the rhyme, or the number of moras, when they nominate a syllable as heavy. In Golin, an East New Guinea Highlands dialect of Chimbu (Bunn & Bunn 1970), for instance, all syllables carrying a high tone are heavy for stress. Since there is no actual reference to a quantitative or segmental difference between “heavy” and “light” syllables, it is perhaps confusing to speak of weight in this case. It is not even possible to shift the explanatory burden to sonority (which we may do in other languages, cf. Zec 1988, or the discussion on Inga below), because we would not want to attribute high sonority to high pitch. The existence of languages such as Golin, in which stress rules are sensitive to other syllabic properties than sonority or weight, forces upon us the view that the perceptual salience represented by the peaks in figure 9 from chapter 4 (which we repeat below for convenience) does not only consist of moraicity and sonority, but something more, to which stress rules can also refer.

---

[8] Much more could be said about these Count Systems. For instance, what the domain of main stress assignment should be (bounded or unbounded) is unclear. The whole word is used to calculate the position of main stress, yet main stress seems to fall in a disyllabic domain at the word edge. Furthermore, a discussion of Warrgamay four syllable words with an initial long vowel would open the debate on the allowance of uneven trochees (those which have a heavy head and a light dependent) which could also be used for *mú:ba*. This chapter, however, is not intended to discuss such theoretical issues in too much detail.

[9] Analyses of these patterns can be found in Hayes (1995), who also notes problematic cases and expands the theory with some extra devices to handle them.

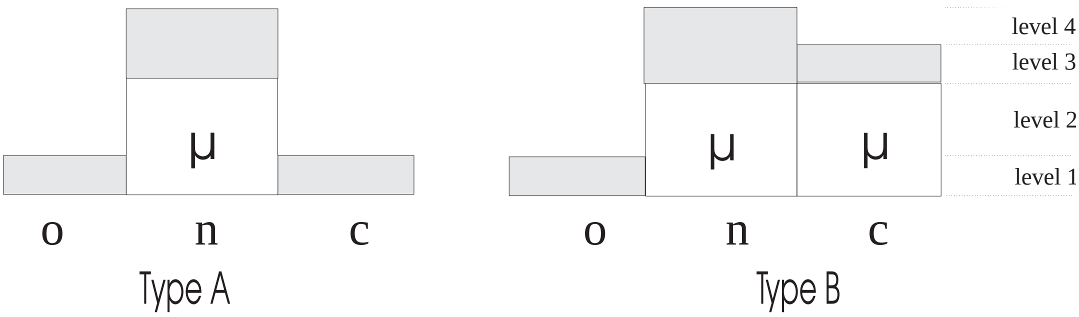


Figure 1: schematic prosodic syllable as relevant for phonological weight (dark areas indicate prominence added by prosodic features).

We have called the complete set of potential stress-influencing properties *prominence* (after Hayes 1995) and assume that the main contribution to prominence in Golin is made by tone. From here on we will call systems like that of Golin prominence systems. It is not difficult to understand why Golin high-toned syllables are prominent. High tones are perceptually more salient than low tones or level tones. We do not need to refer to syllabic structure or mora numbers to describe the Golin prominence differences. It is merely a question of the presence or absence of a high tone. The prominence differences at level 4 depend on the height of the tone on the vowel, the higher the tone the more prominent the syllable. The barrier that separates the high tone "heavy" from the low tone "light" syllables lies somewhere in level 4, and the metrical rules can refer to this difference. These rules stress the last "heavy" syllable in a word, or, in case there are none, the last "light" one. This means that the feet we need are unbounded, and that the default unbounded foot will be right-headed. Some examples are given in (21) (á denotes high tone, x a prominent foot head).

(21)	(. . x) ó wá ré	'bat'	(. x .) sí bági	'sweet potato type'
	(x . .) ákola	'wild fig tree'	(. . *) kawligi	'post'

The structures in (21) reveal an important fact about the way in which we view stress assignment in Golin. It does not seem to be the case that prominence-sensitive rhythmic feet are built to parse the word, and that the last head is promoted to main stress by an End Rule. It is more likely that there is no rhythmic level at all, and that a prominence-sensitive End Rule "scans" the word for high tones and assigns main stress to the last one, assigning a default stress to the very last syllable if no high tones can be found (such a rule could best be described in the Idsardian formalism, cf. chapter 1 footnote 9, but, as mentioned above, we will not choose a particular formalism in this chapter). This side we choose in a possibly very thorny theoretical debate is not crucial to the arguments presented in this study, however. What matters is the fact that the Golin stress rule is sensitive to tone. At the end of chapter 4 we postulated that only those syllabic/segmental properties that are relevant to prosody can potentially influence stress placement. According to these limitations we may include tone in the set of prominence features since it is clearly a prosodic phenomenon.

Tone is not the only feature that can add to syllabic prominence.[10] Remember that we introduced *sonority* as the most important prominence component besides moraicity in chapter 4. It is well known from the literature that some languages have stress rules that refer to sonority. In Inga (Quechumaran-Colombian, cf. Levinsohn 1976), for instance, syllables that are closed by a sonorant consonant (m, n, r or y) are heavy (Inga does not have long vowels). If the final syllable is heavy it carries main stress, in all other cases main stress is penultimate.[11] Examples are given in (22).

(22)	yawár	'blood'	kánčis	'seven'
	apamúy	'to bring'	kamkúna	'you (pl.)'

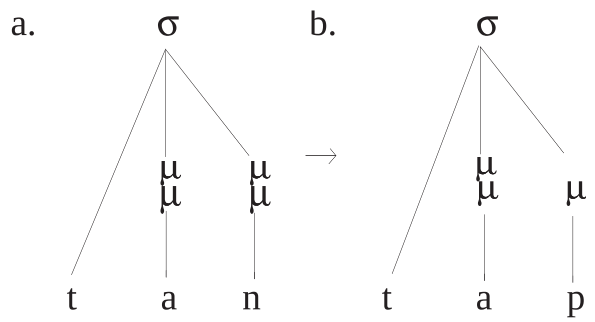
We presume that usage of the word 'heavy' is again out of place here. There is no reason to believe that considerations other than sonority of the coda consonant feed the stress rule. As an alternative to a sonority-sensitive Weight-by-Position rule, we might just as well assume that the WBP is active to begin with, and that the difference between "heavy" and "light" syllables is made at level 3 (cf. figure 1). Syllables with sonorant codas exceed the prominence threshold in two segments while those with non-sonorant codas only do so in the vowel. It seems, at least, that there is no quantitative difference between the heavy and the light CVC syllables in Inga. We can think of no reason why the difference between, for example /n/ and /p/ should be moraic, as if the difference were based

[10] It is not our intention to give a complete listing of the kinds of feature that can contribute to prominence. We merely note the existence of prominence and that it is likely that all the features that contribute to it are prosodic.

[11] See also Zec (1988) for a discussion of variance in the weight/prominence of coda consonants in other languages, and Kenstowicz (1994) for a discussion of languages in which vocalic sonority influences stress.

on the presence or absence of a segment rather than on sonority. These considerations withhold us from following Hayes in adopting a separate moraic grid for these cases. He claims that languages may exploit differences in mora structures as those depicted in (23).

(23)



The difference between Inga syllables with sonorant coda consonants and those with non-sonorant codas, is thus expressed as a quantitative difference. As we have claimed above, we do not believe that quantity plays a role in languages like Inga. We suspect that all prosodically active properties that can interfere with stress assignment must be represented on the non-moraic levels in figure 1. Therefore, we would rather group Inga with languages like Puluwat (cf. Elbert 1972, and section 5.3.2) in which moraicity plays no important role either, but prominence does. In this language syllables *beginning* with /h/ are stressed. It would be very unwise to represent this as weight on a moraic grid. We seem to need prominence level 1 that we introduced in chapter 4 here and assume that the Puluwat “heavy-light” distinctions are made on this level. We can imagine the prominence of /h/'s to differ from that of other onset consonants (see 5.3.2.1 for discussion). If we adopt the role of prominence levels in Inga as well as in Puluwat, we do not need Hayes’ moraic grid.

Admittedly, there might be languages in which an opposition between CVV, CVS and CVO exists (S is sonorant, O is obstruent), as opposed to Inga CVS vs CVO. In these cases the difference between (23a) and (23b) would nicely show the difference between CVV and CVS. CVO would then lack the final mora entirely. However, we believe that some types of mixed systems may exist that combine prominence (like that in Inga) and weight proper. Therefore, we will not adopt the moraic grid, but analyse languages that have a CVV, CVS, CVO-like opposition as complex prominence/weight systems, or in our own terminology: as systems that operate at two prominence levels, one of which is level 2. In section 5.3.2 we will describe the prominence/weight interaction of Pirahã, in which not the sonority of the coda, but that of the onset seems to determine syllabic prominence. In that section we will only deal in some detail with other languages like Pirahã, since these are much more relevant to the main topic of this thesis than coda-prominence languages like Inga. First, however, we will introduce a group of languages that seem to have moraic onsets, but which, after closer inspection, appear to be neither quantity- nor prominence-sensitive.

5.3 Onsets as a factor in stress rules

In the previous chapters we have provided evidence for the claim that onsets cannot contribute to syllable weight. As was noted in section 5.1, though, some of the languages we find in the literature do seem to refer to onsets when determining the placement of main stress. Many of these languages seem to be straightforwardly sensitive to “onset weight”.[12] Australian Cape York languages like Linngithig, Mbabaram, Kuku-Thaypan and Lamalama, the Australian (Arandic) languages; Alyawarra, Kaytetj and Western Aranda, and the American indian languages Banawá (Amazonian) and Iowa-Oto (Siouan) all have the following stress rule: “stress the first syllable that begins with a consonant”.[13] It has been argued in the past that these languages provide evidence for the existence of onset weight (cf. Davis 1985). However, we have argued extensively that moraic onsets do not exist. Besides, if such a straightforward weight distinction really existed, we would expect many more languages to employ it. Our view is further supported by the fact that all of the languages mentioned above can easily be captured in an analysis that does not refer to any sensitivity whatsoever. Let us now then show why these regular “onset-sensitive” languages are not truly quantity- (or prominence-) sensitive.

[12] We will only deal with onset-sensitive stress assignment rules. For a treatment of other phonological phenomena that seem to involve onset weight see, among others, Downing (1998) and Pensalfini (1996). See also Davis (1988) for marginal onset phenomena in Italian and English and Dainora & Steinberg (1996) for such phenomena in Canadian French.

[13] A list of Aboriginal languages for which we found this stress rule can be found in VIII in appendix D. The same appendix (in IX) contains references to some of the other exceptional Aboriginal languages mentioned in this section.

5.3.1 Onset weight and Initial Dropping

Some of the best documented languages in the group mentioned above are Western Aranda (Strehlow 1942), Alyawarra (Yallop 1977) and Banawá (Buller, Buller & Everett 1993). Western Aranda and Alyawarra are closely related languages that have the same stress rule for words of more than two syllables. This stress rule has already been mentioned above. As we can see in (24), the first syllable that has an onset carries main stress, while every other syllable thereafter receives a secondary stress. The stress rule for Banawá is much more complicated, but if we consider the Banawá words in (24), we observe that it is similar to Aranda and Alyawarra in that it puts some degree of stress on the first syllable that has an onset.

(24) *Western Aranda*

káːputa	'head'	ibátja	'milk'
wóratàra	'place name'	arálkama	'to yawn'
lélantìnama	'to walk along'	ulámbulàmba	'water-fowl'

Alyawarra

párriyka	'fence'	ilípa	'axe'
ráthirra	'they two'	apmpírnitjìka	'will cook'

Banawá

tátikùne	'hair'	ufábunè	'I drink'
mákarì	'cloth'	atìkadámuèi	'forget'
tìasíanì	'acquire'	uwárei	'make noise'

These data can easily be misinterpreted as evidence for onset-sensitivity. Davis (1985) describes Aranda as a quantity-sensitive system. He claims that Aranda onsets contribute to syllable weight and assigns a mora to the onset and the nucleus (and attaches the coda directly to the syllable node), dividing the syllables into a light V(V)(C), and a heavy CV(V)(C) set. His Aranda stress rule then becomes: stress the first syllable if it is heavy (i.e. bimoraic), otherwise the second. This yields correct surface patterns for the words in (24). Within contemporary metrical theory, however, we cannot accept this analysis. As we have argued above, onset weight does not seem to have any phonological or phonetic reality. Apart from the criticism on the possibility of moraic onsets, we note some other serious drawbacks of analyses for Aranda stress that recognise onset weight, like Davis'. Even if we accepted the possibility of onset weight, we fail to see how we could then leave out the weight contribution of the coda. In Davis' analysis of Aranda, codas are non-moraic. We reject this possibility in view of the evidence presented in this dissertation, which shows that codas are better candidates for moraicity than onsets, and hence, languages would choose to work with moraic codas sooner than moraic onsets (if the usage of moraic onsets were an option). If we keep to the rule that codas must add to weight if onsets do (rule 3 in (4), chapter 4) we must conclude that the first syllable in words like *ergúma* 'to seize' is as heavy as the second.[14] Hence, we wrongly derive stress on the first syllable because it is the first heavy syllable of the word.

Moreover, Davis' analysis misses some crucial observations. Second syllables, or for that matter, any syllable other than the first, almost always have an onset (for an explanation see the discussion on ID below). Hence, only initial syllables can be light in Davis' analysis, and weight differences between syllables other than the first and the second cannot be based on onset presence. In fact, rhythmic assignment of feet to the right of the first foot does not respect weight at all in Aranda (even in the rare cases where onsetless syllables are not initial).

We hope to have shown that a quantity-sensitive analysis of the Aranda-type stress rule is untenable. The facts direct us towards a quantity-insensitive analysis of Aranda stress in which some other mechanism accounts for the observed differences in main stress position. To find a clue to what this mechanism might be we must make a brief digression into the history of the Aboriginal languages in question.

Below we present some of the other languages that have stress patterns like Aranda, Alyawarra and Banawá. With the exception of Mbabaram, these languages are all located on the Cape York peninsula in Australia. For these Cape York languages, an explanation for the rare stress placements is readily available. We claim that this explanation at least applies to all the Australian languages mentioned in appendix D, VIII.

Dixon (1970, 1991) writes about Mbabaram (a language from the Cairns rainforest, near Cape York) that it usually stresses the first syllable that begins with a consonant, as in (25). Mbabaram does not have many trisyllabic words, but the pattern becomes clear from a comparison of some disyllabic words.

(25)	wánu	'termite'	aŋál	'boomerang'
	búmba	'ashes'	albí	'grey'

Many Cape York languages show the same characteristics as the

[14] We owe this example to an anonymous reviewer for The Linguistic Review.

languages in (24) and Mbabaram. We exemplify this in (26).[15]

(26) *Lamalama* (Laycock 1970)

tútulu	'black'	arkúlan	'moon'

Umbuygamu (=Morrobalama, Ogilvie 1994)

túfu	'cabbage tree'	adár	'tongue'
gápar	'belly'	alílir	'carpet snake'

Agwamin

máta	'stone'	abó	'ground'
kérindja	'star'	arábin	'moon'

In Dixon (1991), among others, we find a possible explanation for the metrically unorthodox behaviour of this group of languages. He claims that Mbabaram underwent a process in which initial CV was dropped in all words and, in case the initial syllable had a long vowel (CVː), the remaining second V was replaced by /a/. We illustrate this Initial Dropping (ID) phenomenon in (27).

(27)

gúyu	→	yú	'fish'
bámba	→	mbá	'belly'
yíːbar	→	abér	'south'
wáːŋal	→	aŋál	'boomerang'

It appears that almost all of the relevant Cape York languages also have an ID phase in their history, though not all of them replaced initial long vowels in initial syllables by /a/. In most cases the remaining short vowel kept the identity of the long vowel.[16] Alpher (1976) mentions Aranda as an ID language in central Australia. An important result of ID was that onsetless syllables were created word-initially in languages that otherwise only had syllables with *obligatory* onsets.

Dixon proposes that Initial Dropping is the result of a prior shift of stress to the second syllable. The initial syllable would then have become subject to stress conditioned lenition. This explanation would hold for Lamalama and the other northern Paman languages, but other languages, like Olgolo (Dixon 1982) have dropped the first consonant without losing

---

[15] We present Lamalama as the representative of a whole group of languages. See Laycock (1970) for the members of this group.

[16] For an extensive, theoretically oriented, overview see Smith (1997).

following short vowels or shifting stress, as is illustrated in (28).

(28)	báma	→	áma	‘man’
	gúda	→	úda	‘dog’
	míṉa	→	íṉa	‘animal’

This means that the loss of initial consonants in these languages cannot have been conditioned by stress shift, which might indicate that stress shift is the result of ID rather than the cause. Had stress shift been the cause of ID we might expect to find languages in which the stress has shifted to the second syllable, while the initial consonant has not yet dropped. As far as we know such languages do not occur in the areas where ID has taken place, while the converse case is exemplified by Olgolo.[17] In any case, after the loss of the initial consonant, all Olgolo words began with a (stressed) vowel. Dixon makes clear that having stress on initial vowels (the unique onsetless syllables) makes the language highly unstable, and shows that Olgolo speakers do attempt to resolve that situation. Occasionally, his informant spontaneously uttered the post-ID forms in (28) with either /w/, /y/ or /ṉ/ added in initial position. Note that, in languages like Aranda, the stress is not offending since it does not occur on word-initial vowels. We predict that such languages are less inclined towards reintroduction of initial consonants. Such reappearance has indeed not been reported for languages of the Aranda type.

From the ID facts we might conclude that we do not need onset-sensitivity to account for the stress rules of the Cape York languages. The result of ID, however it came to be, is that words which had a CVCVCV pattern in their pre-ID form will have a CVCV pattern in their post-ID form (underlining denotes stress). Likewise, words with initial long vowels went from CVːCVCV to VCVCV. Stress moves to the second syllable in both cases, but after ID that movement is only visible in the second case. We are left with a stress system that seems to count the onset when determining the weight of a syllable, while we can easily deduce from the history of the languages that weight does not have to play a role in stress assignment.

We do not claim that ID is the solution to our problem, it merely points in the direction of one. We presume that it provides at least some

[17] Mbabaram has some words that show stress on the second syllable in spite of the fact that they begin with a consonant. These do not form a counter-argument, however, since these words have also lost their initial syllable. Clearly some other process yields these deviant stresses.

evidence for the claim that the languages mentioned above can be reanalysed through ordinary stress rules that are hampered by edge effects. It will be clear that the ID rule “mutilated” the initial syllable to a large degree. In an analysis that has to account for the positioning of stress in languages like Aranda, we might assume that this initial syllable is therefore defective and unable to bear stress. The result would then be that the stress shifts off the first syllable if it is deleted entirely by ID (which is logical), but also if the first syllable is shortened from CVV to V. In this way we can analyse the languages in question without reference to quantity, as was the case before ID took place, while respecting the observation that “onset-sensitivity” only occurs at the beginning of words. The exact technical details of the rules we need to keep stress from being placed on defective initial syllables will be discussed in section 6.2, where we present a metrical analysis for Western Aranda and Alyawarra. Another strategy that would serve to avoid stress on these defective first syllables is to reintroduce the initial consonants, as we have seen in Olgolo.

We do admit that we have no reason to handle Banawá and Iowa-Oto (Robinson 1975) in this way. It is probably too far-fetched to assume that these languages have undergone ID. We suspect, however, that comparable processes in their history have yielded the similarities with the Australian languages. We will not include Banawá and Iowa-Oto in the discussion in chapter 6, but Everett (1995) provides a treatment of the Banawá facts that does not involve weight or prominence in onsets. He does not, however, present the kind of historical context that we have been able to give for the other languages, through reference to ID. Whether there was a stage in the development of Banawá in which vowel-initial words did not exist, and an ID-like phenomenon changed that (with the familiar results), remains an open question. We assume that the same holds true for Iowa-Oto.

We have shown here that languages, in which the stress rule seems to differentiate between V-initial and C-initial syllables, need not be analysed in a QS fashion. Diachronic (ID) evidence points in the direction of an analysis that respects the isolated occurrence of onsetless “light” syllables. We have adopted the view that these syllables are anomalies that have to be rendered extraprosodic in some way. Such views pave the way for straightforward QI analysis of all the languages that share this stress rule. The analysis that will be presented in section 6.2 for Aranda and Alyawarra may be applicable to the whole group.

We conclude that we do not need onset weight for any language we know.[18] Remember, however, that in section 5.2.3 we mentioned the existence of languages for which we did seem to need prominence. In the next section we will describe the more complicated cases, in which onsets contribute to syllabic prominence.

5.3.2 Onset prominence

In the previous section we have seen that we can shorten the list of the so-called onset-sensitive languages drastically. Most of the languages that were on the list have the same stress rule as the languages discussed there. So we might claim that all the languages on the list are insensitive to onset weight, as well as onset prominence. There are, however, some languages in which onsets do seem to play a role. This may happen in languages that make the distinction between heavy and light syllables on prominence level 1 in figure 1. Below we will see two such cases, one that is quite straightforward, and one that seems to involve both prominence and weight. We will close this chapter with the introduction of some languages that also seem to employ onset prominence. However, in a prominence analysis these languages would need to refer to prosodically inactive features. We have excluded that possibility in section 4.5, which means we must look for alternatives.

---

[18] In Eastern Popoloca (Kalstrom & Pike 1968) and Trukese (Churchyard 1991), though, geminate onsets are claimed to influence stress. For these languages QS analyses would be needed. Since the contemporary structural representation of geminates involves a moraic link to the second moraic position of the preceding syllable, however, we assume that these systems are more ordinary than they seem at first sight. Suggestive evidence for this view comes from Popoloca, in which length can be expressed either as gemination of the final onset, as in *coyet:ó* 'will get fat', or as lengthening of the prefinal vowel, as in *kočí:ka* 'chicken'. The stresses seem to fall on the syllable that is lengthened. In an alternative view, this hints at a lengthening process that targets the prefinal syllable and assigns to it an extra mora. This mora can then be filled with either vocalic material from the nucleus of the target syllable, or by consonantal material from the onset of the following syllable, rendering this onset a geminate. The choice between the two may be made on the basis of stress. The assignment of these stresses may not be based on QS rules, however. We find many instances of lexical stress assignment in Popoloca, as evidenced by pairs like *thá:ko* 'is teaching' and *thak:ó* 'early in the evening'. So, variation in the location of stress may be due to lexical marking. Therefore, we might just as well assume that stress is assigned lexically to the penult in a small group of exceptions, and default to the final syllable in the rest of the words, and that lengthening chooses either the vowel lengthening or the gemination option, depending on the presence of a lexical stress mark on the penult. Something similar might be going on in Trukese.

5.3.2.1 Puluwat

One of the dialects of Carolinian, an Austronesian language of the Micronesian group, is Puluwat. By the looks of it, this language has a very simple stress rule. Elbert (1972) notes that “CVCV words seem stressed on final vowels”, and that “syllables beginning with /h/ are stressed”. From the examples given in (29) we may conclude that h-initial syllables can at least draw stress onto themselves.[19]

(29)	kiyó	‘outrigger boom’	yiwé	‘when’
	pahálo	‘to drift away’	yiwá	‘where’
	yapwaháno	‘to dry out’		

A provisional conclusion we may draw from the examples we have, is that the Puluwat stress rule stresses the penult if it is h-initial or, in case it is not, the last. Whether this conclusion is right or not, it appears at least that we have to do with a stress rule that can distinguish between h-initial and other syllables. Notice that this case is different from the cases in the previous section, because here the supposed onset sensitivity is not restricted to one of the edges. It seems we will have to deal with Puluwat in the fashion in which we dealt with Golin. We would not want to assume that onsets contribute to syllable weight, let alone say that only /h/’s can do so. It is more plausible to claim that the prominence of /h/ is very low compared to that of the other consonants (cf. Dogil & Luschuetzky 1990). In our framework the h-initial syllables in Puluwat would not exceed the threshold at level 1, while the others would. Note that, in these h-initial syllables, the *difference* between the prominence of the vowel and the initial consonant is the largest. Suppose we would assume that the salience of the transition between the onset and the nucleus is what matters in cases like Puluwat. We could measure this salience with respect to level 1; all segments that remain below the threshold are low enough in prominence to warrant a large enough prominence transition. In Puluwat, only /h/ would do so. Then we can analyse the Puluwat examples using unbounded feet (sensitive to level 1 prominence) and a right-headed default. In many respects the structures presented in (30) resemble those presented in (21) for Golin.

---

[19] There are not enough examples to draw conclusions about possible settings of the stress parameters. The description hints at an unbounded system with onset prominence, but from the examples we might just as well conclude that stress is penultimate except in disyllabic words. Unfortunately, conclusive examples are lacking. We assume, therefore, that the source is right, and that /h/-initial syllables attract stress. Tone is not indicated here.

(30)	(. x .)	(. . x .)	(. *)	(. *)
	pahálo	yapwaháno	kiyó	yiwé

The nature of the stress assignment rule in Puluwat shows the need for prominence level 1 in figure 1. The number of cases for which we need it though, is extremely small, as we predicted in section 4.5. The only other case we know for which we would need onset prominence is Pirahã, a language in which onset prominence seems to interact with normal quantitative (level 2) weight.[20] We must also admit that the Puluwat case is not very strong. We do not have the examples that we need to draw firm conclusions.

However, if we do indeed need prominence level 1, then we must deal with the one striking property that sets it apart from the other levels; level 1 prominence is calculated with respect to the prominence of the nucleus (so that onsets with a low sonority seem to be the most prominent ones), exactly the reverse of what we do with the other levels. This is not just a peculiarity of Puluwat. We will see it again in the discussion of Pirahã in the next section.

5.3.2.2 Pirahã

Pirahã (Everett & Everett 1984) is a Mura language spoken in Brazil. At first sight the stress rule in Pirahã seems to be sensitive to both the presence and the identity of the onset, as well as to vowel length. Remember that we defend the claim here that the mere presence of an onset can never be the basis for a metrical heavy/light distinction. Therefore, we prefer metrical reanalyses of the Pirahã stress rule that are not sensitive to onset weight, like those of Everett (1988), Davis (1988) and Hayes (1995). These analyses do feature onset prominence, which is a characteristic of the Pirahã stress rule that we cannot circumvent. As they stand, however, these analyses make use of the prominence device in a way that cannot easily be reconciled with the views on prominence that we have defended in the previous sections. In this section we will present an analysis that is in the spirit of the earlier analyses mentioned above, and state what seems to be wrong with it if we keep to the views

---

[20] Another language in which this happens on a more modest scale is Asheninca Campa (Payne 1990). In this language four prominence categories are distinguished: CVV > Ca,Ce,Co,CiN > Ci > si,ši. Hence, the sonority of the onset only serves to separate the last two categories. Hayes (1995) even deletes the last category from the scale altogether, attributing the effect to segmental rules. The remaining stress patterns are interesting but they do not feature onset prominence. Onset-insensitive analyses of the Asheninca stress pattern may be found in Hayes (1995) and Goedemans (1996a).

schematised in figure 1. We will conclude with an alternative that turns the Pirahã prominence rule around 180 degrees, bringing it in line with more general prominence-sensitive systems.

Let us first look at some Pirahã data, taken from Everett (1988). (Tone is not indicated, syllables are separated by a "." for convenience.)

(31) a. káː.gai 'word'
b. ho.aa.gái 'species of fruit'
c. ʔá.ba.gi 'toucan'
d. ʔa.ba.pá 'Amapa (city)'
e. kao.bii.ga.bái 'almost falling'
f. pia.hao.gi.so.ái.pi 'cooking banana'
g. pii.kao.bíi.ga.ha 'was certainly raining'

Everett notes that stress is assigned to the rightmost "strongest" syllable in a three syllable window at the end of the word. Pirahã does not have secondary stress. We can see that the onset is relevant in the determination of Everett's "strength". In (31a) there are two heavy syllables, of which the one with the voiceless onset receives main stress. In (31b) the heavy syllable with the onset bears main stress, and not the one that has no onset. (31c and d) show that, if the vowel is short, the voiceless onset is also stronger than the voiced one, and that /ʔ/ patterns with the voiceless consonants. Of the two equally prominent syllables /bii/ and /bai/ in (31e) the rightmost is stressed. From the example in (31e) we may also conclude that stress can indeed only occur on one of the last three syllables. In this example a syllable that has a voiceless onset and a long vowel occurs to the left of three syllables that have voiced onsets and it does not receive stress. (31f and g) reveal that syllables with long vowels are always stronger than syllables with short vowels, irrespective of onset voicing. So, calculation of the strongest syllable proceeds through a comparison of the weight and prominence of the candidates. According to Everett, prominence lending characteristics of the syllable are 1) the mere presence of an onset, and 2) voicelessness of this onset. A scalar representation of Pirahã prominence/weight distinctions shows that prominence and weight interact both in heavy and light syllables.

(32) Pirahã Weight Scale (G = voiced consonant, C = voiceless):
CVV > GVV > VV > CV > GV

Since stress can be assigned to any one of the last three syllables, we cannot immediately decide whether the system is bounded or unbounded.

For a theory that normally only allows binary or unbounded domains such three syllable windows are a problem. Hayes (1995) states that the last syllable must be marked for EM in order to derive stress on the antepenultimate syllable, but he also needs to be able to stress the final syllable. To keep this possibility open he incorporates the extrametrical last syllable into the final foot in a fashion not unlike *stray syllable adjunction*, known from older versions of metrical theory. He justifies this move with the help of an observation made by Everett (1988). He notes that the 'real' EM nominalising suffix *-sai* (which never receives stress though it is always the rightmost strongest in the word) only allows a **two** syllable window to the left of it, in which main stress is assigned:

(33) (. *) (* . .)
ʔi.to.pi.<sai> 'remover' not possible is: *ʔoi.boi.bii.<sai>

Hayes assumes that *-sai* is marked for not undergoing the adjunction phase. This example constitutes strong evidence for the derivation of ternary domains via EM instead of deriving them directly. Alternatively, we could assign stress in a binary domain at the right word edge and shift stress out of that domain if the syllable to the left of it is heavy with respect to the syllables in the domain.

To select the position for main stress in the domain we employ the hierarchy in (32) to assign grid columns of different height. All the authors mentioned above use these grid columns, but Hayes is the only one who explicitly mentions the fact that the grid columns assigned to the Pirahã syllables are prominence based. He assigns *'s to syllables on a prominence plane according to the following scheme:

(34) CVV: *****
GVV: ****
VV : ***
CV : **
GV : *

This information is then accessed by an End Rule which places stress on the rightmost highest grid column. Because Pirahã has no secondary stress we assume that no footing takes place prior to application of the End Rule, as in Golin. Hence we build one non-iterative foot at the right edge in the manner described above (with EM). A prominence-sensitive End Rule (Right) then assigns stress in this domain. If we do so, we can easily derive main stress. In (35) some examples are presented with their prominence grids.

```
(35)  ( x   .)      ( .   .   x)               ( .  x    .)         ( x   .   .)
      káːgai        ʔa.ba.pá      pia.hao.gi.so.ái.pi   pii.kao.bíi.ga.ha
        *   *         *   *   *                 * *   *              *   *   *
        *   *         *       *                 * *   *              *       *
        *   *                                     *                  *
        *   *                                                        *
        *
```

In view of what we have said previously this analysis has some serious drawbacks. We feel that this is just an indirect way of accepting onset weight after all. A grid mark is assigned on the basis of the presence of a segment, in the same vein as the WBP assigns a mora to codas when present. We reject this possibility. Dividing heavy and light syllables on the basis of presence or absence of segments is a level 2 (weight) operation which cannot be applied to onsets. Prominence distinctions can only be based on non-structural syllabic properties. Moreover, the prominence scale in (32) is clearly the reverse of what we would expect. Sonority is obviously the determining factor in the prominence relations between Pirahã syllables. Yet, the fact that voiceless onsets attribute more prominence to the syllable than voiced ones is surprising to say the least.

In consideration of the criticism voiced above we propose to look at Pirahã from another angle. Suppose that the onset does not directly contribute to syllabic prominence, but that Pirahã is special in that it calculates the prominence of the vowels *relative to that of the consonantal segments in the syllable*, which happen to be onsets in all cases (the language has no codas).[21] The weight distinctions in Pirahã would have to be made at level 2. As usual, long vowels are heavier than short vowels because they exceed the level 2 threshold longer. In addition there is a threshold at level 1. The prominence factor that can make vowels stand out from others is not calculated by absolute vocalic sonority, but with respect to this level 1 threshold. Like in Puluwat, it is the prominence difference between the vowel and the onset that matters. The greater the difference, the larger the relative rise in sonority (and perhaps the degree of spectral change, i.e. PIVOT), and the more salient the syllable. Voice is the prime distinguishing feature in Pirahã. This is allowed in our model since voice contributes to sonority. Voiceless onsets

[21] In this way we introduce relativity in the prominence relations. This reminds us of the way in which we look upon the head and the dependent of a foot. The head of a foot is not a head in an absolute sense, but relative to the weak node in the same foot. In the same fashion we believe Pirahã vowels to be prominent syllabic heads that can only express their prominence with respect to weaker elements in the same syllable.

stay under the threshold, *so the sonority difference between them and the vowel is significant*. Voiced onsets exceed the threshold, and hence, remain too close in sonority to the vowel. An extra advantage of this way of looking at Pirahã prominence is that we indirectly explain why the presence of an onset matters. This presence does not add to weight nor to prominence. It merely opens the possibility to determine the relative sonority of the vowel. Onsetless syllables are less prominent than syllables that do have an onset because the relative prominence of the vowel cannot be determined in the absence of other segments in the syllable. In those cases only the weight criteria apply. Even if we were to adopt the view that relative prominence needs to be calculated for all the vowels, we would be able to maintain this hypothesis. Glottal stops are phonemic in Pirahã, so they would have been indicated had they been the segments that occur between the last vowel of a given syllable and the first vowel of a following onsetless syllable. Since that is not the case, we assume that the gap between the two vowels in question is filled by a glide on the phonetic level. That would mean that the relative sonority of the vowel in the second syllable is extremely low, since glides are the most sonorous of all consonants.

The representations for the analyses of Pirahã words do not alter as much as the ideas behind them. In the analyses of the other prominence-sensitive languages that were presented above, we have not used a prominence grid, since the distinctions made were straightforward enough to represent without such an aid. We could employ it though, if the prominence scale is more complicated, but we would merely use it as a representational device reflecting the distinctions that can be made on the levels of our prosodic syllable. The grid itself is devoid of any theoretical meaning. So, translating the levels and thresholds representation of the Pirahã prominence scale into grid columns we may follow the convention in (34) as well as any other. We assign three *'s to a long vowel, none to a short vowel, add one * if vocalic sonority noticeably rises with respect to the consonant in the same syllable, and one more if this rise exceeds the prominence threshold. Thus, we arrive at exactly the same representations as those in (35), the difference being the reasoning behind them.

This concludes our discussion of the Pirahã data. We have seen that these data can be handled easily if we recognise the role of prominence in the assignment of stress. Our contribution with respect to this analysis has been the realisation that not the prominence of the onset but the relative prominence of the vowel with respect to that onset is the determining factor for the heavy-light distinctions in Pirahã. Dogil (p.c.) notes that the way in which Puluwat and Pirahã use the difference in prominence between the nucleus and the onset (or the strength of the transition between them) as a measure for the prominence of the syllable (and hence, as a stress attracting property) forms fiarly direct evidence for the role of the PIVOT (cf. sections 4.1, 4.3.2 and 4.4) in the perception of syllable weight. Indeed, we have not discarded the possibility that the PIVOT is the trigger for proper duration perception, and thus, phonological weight. It is not inconceivable that languages can also use the strength of the trigger itself in their determination of syllable weight (or prominence). Puluwat and Pirahã would be examples of such languages. In that case, we would expect other onset-prominence languages that we may discover in the future to show the same behaviour.

5.3.2.3 Djapu, Mathimathi and Ngarigu

The three final cases we will discuss are all Australian Aboriginal languages. In Djapu, a Yolŋu language from Arnhem Land (Morphy 1983), main stress is generally located on the initial syllable. In a handful of exceptions, though, we find stress on the second syllable. The striking thing about these exceptions is that the onset of their second syllable is /d̪/. In (36) we present some examples.

(36)	búyuka	‘fire, firewood’	ban̪d̪ány	‘dry, clear’
	ápaty	‘white person’	bud̪ápthu-N	‘go down and cross’

The Djapu exceptions are too infrequent to draw any firm conclusions with respect to onset-sensitive stress rules. Yet, if the pattern were general, the rule we would need to describe it would be: “stress the second syllable if it starts with a /d̪/, otherwise stress the first”.

More interesting in this respect is Mathimathi (Hercus 1986), a Kulin language from the Upper Murray River area. It has been claimed that Mathimathi stress is placed on the second syllable if the first syllable is light (has no long vowel or closing consonant) and the second begins with a coronal consonant or vowel in hiatus. If the second syllable carries main stress, the first gets a secondary stress. Other secondary stresses occur on alternates after the main stress, but not on the final syllable if it is light. Unlike in Djapu, stress on the second syllable is not exceptional in Mathimathi. There seems to be a rule that applies in most cases. The relevant patterns are illustrated in (37).

(37)	bérbaθa	'to jump'	wìθíwaθa	'to come back'
	géŋginìn	'(your) uncle'	wùríŋgi	'grass'
	gágilàθa	'to go hunting'	dùáŋi	'glider(flying possum)'

Ngarigu, a Yuin language spoken further upstream on the Murray (Hercus 1986) reflects this pattern in a less rigorous way. The words in (38) show that the rule may be similar, but there are many exceptions.

(38)	búlburai	'thunderstorm'	guníring	'useless, silly'
	gúginjalà	'kookaburra'	djarímiŋ	'happy, flash'
	málaŋàn	'girl, daughter'	budálag	'tree-goanna'

Considering these data one might suspect that, in Mathimathi, syllables with a coronal onset are heavy with respect to light initial syllables. Such a rule would be odd in anyone's book. It is difficult to maintain that coronal onsets bear a mora, and it would border on the ridiculous to claim that onsetless syllables (which also attract stress) have moraic onsets. Alternative analyses might use onset prominence to handle the Mathimathi case. Both Davis (1988) and Goedemans (1993, 1996a) try to analyse Mathimathi with some sort of rule that respects the special character of coronal onsets. In view of the ideas presented above, however, recognition of coronal onsets as being prominent is not possible. We maintain that only prosodically active properties can add to syllabic prominence. Hence, place of articulation, which is clearly not a prosodically active feature, cannot play a role in prominence relations affecting stress location. Even when we ignore the problems in introducing onset prominence for onsetless syllables, the prominence solution fails.

As was noted above, stress systems like that of Mathimathi must be reanalysed without recourse to onset-sensitivity whatsoever. This may be possible if we recognise the apparent role of onsets in the stress rule as the result of rules that conspire behind the scenes to generate the observed pattern. Such a set of alternative rules is provided by Gahl (1996). Her solution to the Mathimathi problem is modified and placed in a larger context in Goedemans (1997). Of these two studies only the latter makes reference to Ngarigu. An explanation for the many exceptions we find in Ngarigu is entailed by the analysis of the Mathimathi stress rule that is presented there. In section 6.3 we will recount this analysis.

5.4 Conclusion

In this chapter we have introduced the metrical machinery that is needed to phonologically analyse the stress patterns we find in the languages of the world. We have exposed the rules that are a part of metrical phonology in several (simplified) analyses of Australian Aboriginal languages. Against this background we have discussed some languages in which onsets play a role in the assignment of stress. Most of these languages were Australian. For many of these Australian languages (plus one Amazonian and one Siouan) the reported stress rule is to stress the first syllable that has an onset. Since this rule refers to a structural syllabic property with respect to onsets (viz. their presence or absence) its logical theoretical equivalent would count onsets in its determination of syllable weight. Since we deny that possibility, we looked for an alternative. Historical evidence from the Australian languages in question prompted us to search in the direction of an analysis that is not quantity-sensitive to begin with. Some suggestions were given for a possible reanalysis that will be discussed in detail in chapter 6. In that analysis we presume that the first foot with which we parse the words in these languages is misaligned, starting at the second syllable because the first lacks an onset, and hence, is unable to bear stress.

The discussion of these virtual onset weight cases was followed by the introduction of some individual languages in which the identity of the onset determines its role in the metrical rules of the language. For these languages we need the onset prominence level (no. 1) that was introduced in section 4.5. It appears that only very few languages employ this onset prominence option. We have discussed Puluwat and Pirahã (and referred to Asheninca Campa in which such onset influence might be marginally present). Of these cases Pirahã is the most complicated one. For this language, several prominence analyses have been presented in the past. We have shown that the usage of prominence in these analyses is not compatible with what we view to be the role of prominence in stress rules. We have presented an alternative in which we use the more limited form of prominence that we advocate.

It must be noted that, though amply discussed in the literature, the Pirahã data are not undisputed. First of all we observe that Pirahã words may carry several distinct tones, which might interfere with stress assignment in a way we do not fully understand yet. Surely the fact that voiceless onsets, which can induce a high tone on the following vowel (cf. tonogenesis), are more prominent with respect to the stress rule than voiced onsets *could* indicate an active role of tone at some point in the derivation.

Furthermore, we note that even the stress rule of Pirahã itself is not undisputed. Preliminary reports from other researchers in the Amazon area suggest that Pirahã stress is simply penultimate (Aikhenwald p.c.). If that is true, the only motivation for onset-sensitivity comes from the weakly supported case of Puluwat. We might, even now, conclude that the case for onset-sensitivity does not rest on a very solid foundation. After more research on Puluwat and Pirahã we might have to conclude that onset prominence is just as improbable (or even impossible) as onset weight. It might be the case that the prominence levels in figure 1 depend on moraicity and that segments that cannot be moraic (onsets) cannot feature in prominence rules either.

The fact that we consider the domain of prominence to be limited to the influence that prosodically active features may have on stress placement presents us with a problem if we regard languages like Djapu, Mathimathi and Ngarigu. In these languages a prosodically inactive feature seems to influence stress. We introduce the possibility that the stress rule in such languages is not onset-sensitive at all, but that the supposed influence of onsets on stress is only a surface pattern due to the conspiracy of underlying onset-insensitive rules. As we will see in section 6.3, this is probably the case for Mathimathi and possibly also for Ngarigu. We consider this to be an indication that we are on the right track. A reanalysis for Djapu might also be found in the future. Moreover, the fact that we can indeed find such reanalyses for the languages we predict to be impossible prominence systems supports the view that prominence can only come from prosodically active features like pitch and sonority.

6 Two Case Studies

6.1 Introduction

This chapter contains an in depth discussion of some of the languages that were introduced in chapter 5. In two case studies we will discuss these languages in detail. The first case study concerns the languages Western Aranda and Alyawarra, two members of a group of languages that was presented in section 5.3.1. The stress rule that the members of this group have in common places main stress on the first syllable that has an onset. We have argued in chapter 5 that stress in these languages cannot be analysed with the help of moraic onsets. Instead we have hinted at an alternative that makes use of the ill-formedness of initial onsetless syllables, which results from a historical phenomenon called Initial Dropping (ID). It is this alternative analysis that we will present here.

The metrical analyses of the languages discussed in this chapter will be presented in the framework of Optimality Theory (OT; Prince & Smolensky 1993), or more specifically, the Generalised Alignment extension of that theory (McCarthy & Prince 1993). Since this theory is relatively new, a short introduction of it will precede the analysis of Western Aranda and Alyawarra in section 6.2. The usage of OT for these analyses is not crucial, however. The analyses presented in this chapter can be cast in other metrical frameworks without any great difficulties, as is shown in Goedemans (1996a, 1997).

Our second case study concerns the language Mathimathi. In this language we find another type of onset-sensitive stress rule that we consider to be impossible. For Western Aranda and Alyawarra, onsets bearing weight were the impossible factor. In Mathimathi, however, the problem is not weight related. At first sight stress in Mathimathi seems to be prominence-sensitive in that it stresses the second syllable if its onset is coronal and both the first and the second syllable are light. As we have argued extensively in chapters 4 and 5, we believe that only prosodically relevant properties (unlike the place feature coronal) can be active in prominence-sensitive stress rules. We have seen such rules at work in Puluwat and Pirahã in sections 5.3.2.1 and 5.3.2.2. Since the

Mathimathi stress rule cannot be of that type an alternative analysis is needed. In section 6.3 we present an alternative that is based on historical evidence. The underlying claim will be that a rule that is neither prominence- nor quantity-sensitive causes both the odd distribution of stress and the apparent relation between this distribution and the coronality of the second syllable onset. This analysis will be preceded by the presentation of some phonetic evidence which shows that the Mathimathi stress pattern is indeed as odd as Hercus (1986) reported it to be.

Both case studies in this chapter concern languages that were used in the past as central evidence for the existence of moraic, or at least, prominent onsets. If the analyses we present are acceptable, the case for moraic onsets becomes very weak, while our claim that prominence can only be based on prosodically active features receives extra support.

6.2 A quantity-insensitive analysis of Western Aranda and Alyawarra stress[1]

In chapter 5 we identified the Arandic languages of Central Australia as a subgroup of those languages that have the “stress the first postconsonantal vowel” rule as a result of Initial Dropping. Two of these Arandic languages are Western Aranda (also known as Arrernte) and Alyawarra (or Iliaura). Let us first look at the stress pattern of Aranda. Strehlow (1942) notes that words of two syllables are always stressed on the first syllable:

(1)	gúra	‘bandicoot’	lá:nba	‘armpit’
	éra	‘he, she, it’	ílba	‘ear’

If a word of more than two syllables begins with a consonant, main stress also falls on the first syllable. When such words begin with a vowel, however, the second syllable receives main stress. A weak secondary stress is placed two syllables to the right of the main stress, but only if the recipient is non-final. For convenience we repeat in (2) the Aranda words from (24), section 5.3.1, which exemplify this pattern.

(2)	ká:puta	‘head’	ibátja	‘milk’
	wóratàra	‘place name’	arálkama	‘to yawn’
	lélantìnama	‘to walk along’	ulámbulàmba	‘water-fowl’

[1] This section is a slightly rewritten version of Goedemans (1996b).

In chapter 5 we have already explained why an onset weight analysis of these data is not acceptable. We stated there that the Aranda pattern might better be explained through “skipping” of the initial onsetless syllable.

The most straightforward way in which we can skip the V-initial syllables that remain after ID is through usage of extrametricality (EM). We could just make the first syllable of vowel-initial words extrametrical and assign stresses to the remainder of the words. However, we do not think that EM rules can be applied in this limited fashion. When EM is used it should be used indiscriminately, not just for a subset of words. Therefore, Archangeli (1986) and Halle & Vergnaud (1987) propose to apply EM only to the first segment of all words (except the disyllabic ones). The result of this is that the stress bearer (the vowel) of consonant-initial words remains available for stressing, while the vowel of onsetless initial syllables is rendered extrametrical, and hence, is skipped by the metrical rules. They then assign trochees from left to right, assuming that the rightmost syllable is also extrametrical because it can never bear stress. An End Rule (left) promotes the leftmost foot-head to main stress. In (3) the two different patterns are exemplified.

```
(3)   ( *          )        (   *           )
       (* .)(* .)             ( *    .)(*)
     <l>élantìna<ma>       <u>lámbul àm<ba>
```

This solution is adequate, but it suffers from some drawbacks. First of all we note that it is theoretically undesirable to use extrametricality at both word edges. Such double edged EM is unattested elsewhere. We can solve this problem, however, if we assume that degenerate feet are forbidden (see section 5.2.1). The last syllable of *lélantìnama* then lacks a stress because we may not build a degenerate foot over it, and the last syllable of *ulámbulàmba* is automatically incorporated in the final disyllabic foot.

Furthermore, we are wary of an extrametricality rule that targets two such different segments with such different results, namely no effect at all versus skipping of the entire first syllable. This uneasiness increases if we consider the fact that the analysis in question ignores the ID phenomenon (section 5.3.1) which does seem to be relevant. The motivation for a slightly different EM analysis in Goedemans (1993, 1996a) comes from this ID phenomenon (see below). In Archangeli’s analysis, however, there seems to be no motivation whatsoever to apply EM to the first segment of Aranda words apart from the fact that the analysis works.

We also note that disyllabic words present a problem under this analysis. Stresses for those disyllabic words beginning with a consonant are derived correctly, but when they begin with a vowel, the EM rules render the entire word invisible to the metrical rules. Hayes (1995) calls this the “unstressable word syndrome” and notes that it must be relieved; words may under no circumstance surface without at least some metrical structure. In this case we might choose not to apply segment EM (left) so that the initial syllable can be parsed into a (degenerate) foot and receive stress. The fact that the left-hand EM rule is revoked and not the right-hand one implies that the latter is more strongly enforced somehow than the former. Remember that we suspected the appearance of two EM rules in one language in the first place. How these EM rules should then be ordered, and if that is possible at all, is unclear to us.[2]

Further criticism concerns the fact that Archangeli’s analysis cannot easily handle the data from the related language Alyawarra, which is only different in two minor details and should be captured in an analysis of which the basic parameters are equal to those in Aranda. To fully appreciate this point, one must know more about the Alyawarra data, which are presented below.

Yallop (1977) describes the Alyawarra stress system as follows: disyllabic words can have main stress on the second syllable if they begin with a vowel, else stress falls on the initial syllable (compare the examples in (4) to the corresponding Aranda words).

(4)	athá	‘I (erg.)’	kwíya	‘girl’
	iylpá	‘ear’	kíra	‘meat’

Otherwise stress is like in Aranda (cf. 5a), except that words of more than two syllables beginning with a glide also have main stress on the second syllable (cf. 5b). Yallop further notes that words with an initial consonant are far less common than in Aranda.

---

[2] If we follow our suggestion in this respect, and disallow degenerate feet, the initial syllable of vowel-initial disyllabic words is extrametrical and the second syllable may not have a foot of its own because that foot would be degenerate. The unstressable word syndrome must be relieved. Only in these exceptional cases, the strong demand for some minimal amount of metrical structure might bend the other rules in two ways. Firstly, the extrametricality rule might not apply, leaving the word open for the building of a trochee (results in initial stress, the option chosen in Aranda), or secondly, the ban on degenerate feet might be lifted (results in stress on the second syllable, the option chosen in Alyawarra, see below).

(5) a.	ilípa	'axe'	párriyka	'fence'
	apmpírnitjìka	'will cook'	ngkwárlilànima	place name
b.	walíymparra	'pelican'	yukúntja	'ashes'

We can easily devise an account for Alyawarra stress that is similar to what Archangeli proposes for Aranda. This analysis would have the same drawbacks. Besides, it would wrongly predict stress to occur on the first syllable in *walíymparra* because initial segment extrametricality does not make the vowel of the first syllable invisible to the metrical rules. To derive the correct stress pattern the entire first syllable of w- and y-initial words must be extrametrical. Hence, under Archangeli's analysis the (5b) words are cumbersome exceptions. Let us see if we can find an EM analysis that does better.

One possible analysis might make use of a linking vowel that sometimes appears in front of consonant initial words in Aranda (see Breen & Pensalfini, forthcoming; Goedemans 1993, 1996a). The appearance of this linking vowel is even very frequent in both Aranda and Alyawarra. We could use this observation to postulate an *empty vowel* that occurs before consonant initial words, which is phonetically realised only under certain circumstances.[3] This empty vowel may very well represent an initial empty nucleus that remained after ID deleted initial CV syllables. In words that had an initial CVV syllable before ID (of which C was dropped and VV shortened to V or /a/), the remaining vowel fills the nucleus that is empty in consonant initial words. Thus, phonologically, all words would begin with a nucleus. The second step would then be to apply extrametricality to this nucleus.[4] It is not difficult

---

[3] Alternatively, we could adopt Breen & Pensalfini's (forthcoming) proposal and claim that *all* Aranda syllables are basically VC, at least at some underlying level, removing the special status of the first onsetless syllable. Through this assumption they are able to successfully analyse a host of Aranda morphological phenomena. However, the claim that Aranda has VC but no CV syllables is in direct violation of the rule that CV syllables are the most basic type (allowing VC only when CV is also present), which was long thought to be a universal, and continues to be just that for most phonologists (cf. McCarthy & Prince 1986). It also entails that Aranda must have undergone massive resyllabification after ID, before which it must obviously have been a normal CV language. We will not enter this discussion any further here since it becomes irrelevant to our case once we assume, with Breen & Pensalfini, that stress is assigned only after resyllabification to CV syllables has taken place at a higher level (in which case we end up in the same situation as we are in now).

[4] We could also place stress on the second syllable by parsing it into an iamb with the first (irrespective of whether only the first or all syllables are V-initial). We reject this iambic analysis on the basis of cross-linguistic evidence that points in the direction of a trochaic monopoly on the Australian continent (chapter 5).

to see why a nucleus that is sometimes empty and which, when filled, forms an onsetless syllable (in a strictly CV language), should be extrametrical. So we mark the initial nucleus for EM and skip it in the metrical analysis. In this way we derive the observed stress pattern in a fairly straightforward way. (6) repeats the examples in (3) with their new metrical structure. We use left-to-right trochees and prohibit degenerate feet (N represents an empty nucleus).

```
(6) (     *          )        (    *          )
       ( * .)( * .)              ( *   .)(*   .)
   <N>lélantìnama           <u>lámbul àmba
```

Admittedly, we have not yet solved the problem that we have with the [w] and [j] initial words in Alyawarra. However, the above analysis can easily be saved with the help of some crucial observations made by Strehlow (1942) and Yallop (1977). The solution to the (5b) case lies hidden in some remarks that they make with respect to the nature of the glides in both languages. Strehlow notes that the Aranda [w] is like the [w] in English *water* and that [j] is like the [j] in English *you*. These consonantal properties explain why they pattern with the consonants. The situation in Alyawarra is rather different: “word initial *u*- and *i*- are pronounced *wu*- and *yi*- but the semivowels are in this case part of the phonetic realisation of the vowels” (Yallop 1977:19) and phonologically speaking *wa*- and *yu*- are complex vowels rather than consonant vowel sequences (Yallop 1977:20,28). So, it is justified to treat the glide-vowel sequences as complex vowels which are, like all other initial nuclei, skipped.

Archangeli’s analysis can only be saved through one extra assumption. If we apply segment EM (left), we predict main stress to fall on the first syllable in (5b), unless we allow the extrametricality of this segment (which is a part of the nucleus) to percolate up to make the entire nucleus, and thus the syllable, extrametrical. Alternatively we could state that word initial glide-vowel sequences are monosegmental and therefore affected by segment EM. However, we only deduce from Yallop’s remarks that [wu], [ji], [wa] and [ju] are fully vocalic (i.e. occur in the nucleus), not that they occupy only one segmental slot. In view of these facts, our analysis is less complicated than Archangeli’s and, therefore, to be preferred.

However, we must note that extrametricality on the left hand side of the word is very rare in the languages of the world. Therefore, we consider it wise to be sceptical of analyses that use EM (left). Furthermore, our evidence for the empty nucleus analysis is more substantial than the evidence for either Davis' or Archangeli's analysis, but it is not iron clad. We cannot be sure whether this empty nucleus really exists. We feel, therefore, that the analysis can be improved upon.

The OT account that we will present below is an attempt at such an improvement. It is not based on the assumption of onset weight, empty nuclei, EM (left), or percolation. It does, however, also make use of the observations concerning [w] and [j] that were presented above. We will return to this issue below, but let us first have a brief look at the basic concepts of Optimality Theory.

6.2.1 Optimality Theory

Optimality Theory (OT) is a theory that allows us to view phonological problems from a constraint based angle instead of a rule based, or derivational one. An important claim made by OT is that constraints are ordered with respect to each other and operate on a set of possible output candidates. The constraints themselves make up UG but the ordering of the constraints is language specific. For every given input the function GEN(erator) provides a set of possible candidates which are then evaluated on the basis of the ranked constraints. Output forms may violate constraints, but, the higher the ranking of the constraint, the worse the violation, and the higher the number of offences to a certain constraint, the worse the violation. The candidate that violates the constraints minimally is the optimal output. We might say that the candidate set passes through a filter made up of constraints. The filter is designed such that only one candidate can come out at the other end. This candidate is the selected output. The main assumptions of OT are, thus:

1. Violability: The constraints are violable, but violation is minimal.

2. Ranking: Constraints are ranked on a language-particular basis; the notion of minimal violation (or best satisfaction) is defined in terms of this ranking.

3. Inclusiveness: The candidate analyses, which are evaluated by the constraint hierarchy, are admitted by very general considerations of structural well-formedness.

4. Parallelism: Best satisfaction of the constraint hierarchy is computed over the whole hierarchy and the whole candidate set.

The constraints mentioned above are statements about the well-formedness of surface patterns, they are not surface truths, since constraints may be violated. Best satisfaction is computed along the lines sketched out in example (7). A,B and C are the candidates, while X, Y and Z are constraints (X dominates Y, and Y dominates Z). Suppose that A violates X, that B and C both violate Y, and that C violates Z. This is expressed in a tableau as follows:

7.	X	Y	Z
A	*!		
☞B		*	
C		*	*!

The tableaus are used to visualise the evaluation of the candidates. For economical reasons we usually only select some candidates that resemble the output and help to clarify a point made in the argumentation. These candidates are tabulated in the leftmost column. In the first row the constraints are given in their ranking order. Whenever a candidate violates a constraint a "*" is put in the box corresponding to the candidate and the relevant constraint. The "!" indicates that a certain violation is crucial, it means that this particular violation causes the candidate to lose from the other(s). In (7), candidate B violates the constraints minimally, and is, therefore, the optimal output. This optimal output is indicated by "☞".

From the whole package of constraints that make up UG we will only look at the ones that contain statements about stress patterns. To deal with the phenomena we find in metrical phonology we need some constraints that make sure a string of syllables has a certain metrical structure in the output. In principle, we want every syllable in the string to be incorporated in the metrical structure. The constraint Parse-syllable (PARSE) should take care of that.

• Parse-syllable: all syllables must be parsed by feet.

This constraint can only be operative if we know what feet are and what they look like. Since we are only dealing with languages that have trochaic feet we use the following constraint to define the foot.

• FOOTFORM(Trochaic): a foot is a metrical element expressing rhythm type. It has the following shape: (* .)

Prince & Smolensky adopt a constraint called FT-BIN to ensure that all feet are binary (encompass two syllables or moras).

• FT-BIN: feet are binary at some level of analysis (either μ, or σ)

They claim that this constraint is never violated. From the stress systems of many languages we know, however, that monosyllabic feet must occur (see the discussion on Muriny-pata in section 5.2.1). In (8), for example, feet are assigned from left to right. In words with an odd number of syllables this leads to a "leftover" syllable, as in (8a).

(8)	a.	b. Degenerate Feet	c. Catalexis
	(* .)(* .)	(* .)(* .)(*)	(* .)(* .)(* .)
	σ σ σ σ σ	σ σ σ σ σ	σ σ σ σ σ

If this syllable happens to bear secondary stress it must be parsed, enforcing the adoption of a degenerate foot, thus violating FT-BIN, as in (8b). However, Prince & Smolensky choose to "overparse" here. They use a device called *catalexis* (Kiparsky 1991) to assign a disyllabic trochee to the last syllable, as in (8c). Catalexis is more or less the opposite of extrametricality: a piece of metrical structure is built over the space to the right of the word while there is no segmental material for it to dominate. This is only allowed if it is the dependent, or weak, part of the foot that does not dominate a syllable. This move saves their claim that FT-BIN cannot be violated, but only through the adoption of a new principle (Catalexis). We opt not to enter this discussion here but assume that FT-BIN can be violated.[5] This constraint prohibits the occurrence of monosyllabic feet (which we call degenerate feet) when it is ranked higher than PARSE, but when PARSE is ranked higher than FT-BIN it is more important to parse the leftover syllable than to avoid degenerate feet. Hence the leftover syllable is parsed into a monosyllabic foot.

Finally we need to make sure that the feet are properly mounted to the prosodic word and the string of syllables. For this we use McCarthy & Prince's (1993) ALIGN-FT constraint:

• ALIGN(Ft,L/R,PrWd,L/R)

---

[5] This move simplifies the OT accounts of Aranda and Alyawarra. It must be noted, however, that accounts with non-violable Ft-Bin and catalexis can be just as easily given. The choice for violable Ft-Bin does not affect the crucial parts of the analysis, it is just meant to simplify the presentation somewhat.

This constraint aligns the left or right edges of all feet to the left or right edge of a prosodic word (cf. *Directionality* in section 5.2.1). In our case we will need the left variant of this constraint. Note that it is impossible to align all feet in the word to one edge of a PrWd. Only the first foot can fully satisfy the constraint, all others will violate it. Violation must be minimal, which means that the number of syllables between the edge and the foot must be minimal. In the evaluation, a candidate analysis that skips a syllable when building a foot after the first one has one more syllable between this second foot and the left word edge than a candidate in which the second foot immediately follows the first foot. This extra syllable is the one that makes its candidate owner violate the ALIGN-FT constraint worse than the other candidate. (Later, in section 6.2.2 we will find an example of this in tableau 9b, where the choice between two candidates is determined in this way.) Consequently the other feet are neatly concatenated behind (or before) the first. There are no intervening gaps, since these would take a successive foot further from the edge and thus constitute a worse violation of ALIGN-FT.

We have not yet said very much about the ranking of the constraints. It is without question that FOOTFORM will never be violated in the languages discussed here. In chapter 5 it has been argued that perhaps all Aboriginal languages are purely trochaic. Apart from the unstressed initial onsetless syllable, the languages discussed here exhibit the prototypical Aboriginal stress pattern. So, we will leave FOOTFORM(Trochaic) out of the discussion and assume that it is the highest ranked constraint. Furthermore, we observe that words with an odd number of syllables in Aranda and Alyawarra do not have a final secondary stress (cf. *ká:puta*). This indicates that degenerate feet are forbidden, which means that FT-BIN must be ranked higher than PARSE. We must also assume that PARSE is ranked higher than ALIGN-FT. In that case, it is allowed to build successive feet further from the edge, and violate the ALIGN-FT constraint, in order not to violate PARSE. Had it been the other way around, however, a minimal violation of ALIGN-FT (even of only one syllable) would outweigh a violation of PARSE. This would result in words surfacing with only one foot on the right or left edge, violating PARSE but not ALIGN-FT. As such the ranking order between PARSE and ALIGN-FT mimics the setting of the *Iterativity* parameter (see section 5.2.1). In our case we can see that the words have more than one foot (there are more stresses), so we will have to rank PARSE above ALIGN-FT. This results in the following ranking of the constraints:

• FOOTFORM(Trochaic) >> FT-BIN >> PARSE >> ALIGN(Ft,L,PrWd,L)

Were we to exhaustively describe the stress systems of Aranda and Alyawarra we would need another constraint telling us where the main stress is located. However, this section only deals with the effects of onset sensitivity on the location of stress. Since for this discussion it is not important whether the first stressed position of Aranda and Alyawarra words is a main or secondary stress, we will abstract from it. In the following section we will apply the constraints in the ranking above to the Aranda and Alyawarra facts. We will see that they are not sufficient. Two more constraints are needed, one that has already been proposed by Prince & Smolensky and one that could have been.

6.2.2 An OT analysis of Aranda and Alyawarra stress

Now that we have introduced the theoretical machinery of OT we can try to apply it to Aranda and Alyawarra. Let us begin with the simple Aranda cases that have a first syllable beginning with a consonant. Cases like *wóratàra* and *lélantìnama*. If we cast these in the kind of tableau Prince & Smolensky use to visualise optimality we get (9). As was the case in (7) the first row reflects the ranking order of the constraints. Note that, in the last column, violation is computed in terms of “x syllables from the edge”, hence the deviant symbols. The boxes that are not needed to determine which of the candidates will be the output are shaded.

9a. *woratara*	FT-BIN	PARSE	ALIGN-FT
(* .) wóratara		*!*	
☞ (* .)(* .) wóra tàra			σσ
(* .) worátara		*!*	σ

9b. *lelantinama*	FT-BIN	PARSE	ALIGN-FT
(* .)(* .)(*) lélantìnamà	*!		σσ/σσσσ
☞ (* .)(* .) lélantìnama		*	σσ
(* .) (* .) lélantinàma		*	σσσ!

In (9a) we see that no candidate violates FT-BIN. The first and the third candidate violate PARSE, so they get asterisks. Since violation of a highly ranked constraint carries more weight than violation of one or more lower ones, we can immediately see that the second candidate must win, whatever happens in the ALIGN-FT column (both first *'s in the third column have an !). Since the generator produces every possible option, there are more candidates that could be compared to the ones in (9a). For obvious reasons of space it is accepted practice to only compare the most likely candidates in the tableau. In (9b) the situation is only slightly more complicated. The first candidate violates FT-BIN, which is crucial because the others do not. Both other candidates violate PARSE, so we must look at ALIGN-FT to choose between them. For both candidates the first foot does not violate ALIGN-FT. The second foot of the second candidate occurs two syllables away from the left edge, hence the two σ's in the box. The second foot of the third candidate is located three syllables from the edge, of which the third constitutes the crucial violation. We see here that constraints can be violated gradually, and that the candidate with the worst violation loses.

It may be clear that, if we adopt these constraints in this order, the evaluation of candidates for words beginning with a vowel will go wrong, because the stress pattern is different in these cases. These candidates may not violate PARSE and ALIGN-FT, so a trochee will be built over the first two syllables as in (9), resulting in initial stress, while these words must have stress on the second syllable. We can only solve this dilemma if we somehow allow candidates to violate PARSE and ALIGN-FT. It may very well be that in Aranda and Alyawarra, for some reason, a previously lower ranked constraint has climbed up in the ranking to interfere with stress assignment. A likely candidate is formed by a constraint that has not yet been proposed in OT, but we presume it has been assumed to be present since the beginning. The incentive for the line of reasoning followed below was given in McCarthy & Prince (1993). They claim that their familiar ONSET constraint can be restated in alignment terms as follows:

• ALIGN(σ,L,C,L)

which says that the left edge of the syllable must be aligned to (the left edge of) a consonant. In other words: all syllables must have an onset. McCarthy & Prince note themselves that the number of argument settings is restricted in these cases. It seems that the right-edged variant of this alignment constraint cannot feature consonants, while the constraint above cannot feature vowels. This right-edged variant might be the alignment interpretation of No-coda (the constraint that prohibits the occurrence of syllables with codas):ALIGN(σ,R,V,R). Furthermore, alignments of left edges to right edges (and vice versa) do not occur any more than they do in other fields of the Alignment theory. Therefore we leave out the predictable second edge indicator in the discussions below. From the remaining possible expansions of this alignment constraint, one line is worth exploring in more detail. Suppose that these constraints can also operate on the foot and the prosodic word level. Two possible constraints are then:

• ALIGN(Ft,L,C): every foot must begin with a consonant
• ALIGN(PrWd,L,C): every prosodic word must begin with a consonant[6,7]

The first constraint, which we will call **ALIGN-FTO**, makes sure that the foot is properly aligned to the onset of a syllable, if the syllable has an onset. Thus it is designed to protect syllable integrity.

Downing (1993), Takahashi (1994) and Berry (1996) independently come to the conclusion that the Aranda stress rule is the result of restrictions on syllable well-formedness. Takahashi (1994) and Downing (1998) independently propose constraints similar to ALIGN-FTO. Downing (1998) also mentions a host of phonological phenomena in other languages to which the prosodic expansions of the ONSET constraint are

---

[6] We could continue this discussion using either one of these constraints. Both seem to be applicable to the Aranda and Alyawarra cases. We choose for the foot expansion here, but future investigations may force upon us the second constraint. This typically would be the case if non-initial feet may begin with a vowel. Notice that words like *báuuma* ‘to push’ and *intí:a* ‘cave’ provide evidence for the usage of the foot version of the onset constraint as opposed to the traditional onset constraint. In these (extremely rare) cases, onsetless syllables may be the second, but never the first, syllable of the foot.

[7] Right-hand versions of these constraints are Align(Ft,R,V) and Align(PrWd,R,V). These rules might be connected to phenomena like final consonant extrametricality.

applicable. She analyses these phenomena through interaction of the Onset and alignment constraints, and does not explicitly claim the existence of the foot expansion of the Onset constraint. However, the existence of these phenomena could just as well form independent evidence for the Align-FtO constraint.

Obviously this constraint must be ranked quite low in numerous languages that allow feet to be built on top of syllables that have no onset. But see what happens if we allow this constraint to climb in the ranking and end up above parse. Then it is more important to align the left edge of the foot to an onset than to parse a syllable. This will result in a system in which onsetless syllables are skipped rather than parsed as the *first* syllable of a foot; exactly what we need for Aranda. An evaluation tableau for a vowel initial word is given in (10) (note that the angled brackets do not denote extrametricality here, they simply embrace material that remains unparsed).

10. *ulambulamba*	Align-FtO	Parse	Align-Ft
(* .)(* .) úlambùlamba	*!	*	σσ
☞ (* .)(* .) <u>lámbulàmba		*	σ/σσσ

So far we have not given an evaluation tableau for disyllabic words with initial vowels. We note here that the new constraint must be ranked under Ft-Bin, otherwise words of two syllables of which the first has no onset would be given stress on the second syllable, which is not what we observe. It must be more important to have disyllabic feet than to align the left edge of a foot to an onset. This ordering does not save us completely, because there is an option for disyllabic words to avoid violations of both Ft-Bin and Align-FtO. The word could remain entirely unparsed. Since Parse is the lowest ranked constraint of the relevant three, an unparsed, and therefore unstressed, string would form the optimal output. This problem was noted in the past with respect to other languages and rules were formulated in other frameworks to avoid it (see footnote 2 and the discussion before it). For OT, McCarthy & Prince (1993) have proposed the constraint Lx≈Pr which is defined as:

• Lx≈Pr(*Mcat*): a member of a morphological category *Mcat* correspond to a Prosodic Word.

In other words we may not leave the entire string unparsed, because some structure must be associated to at least part of the word. It will be clear that LX≈PR must be the highest ranked constraint. A disyllabic word like *ílba* can then be derived as in (11).

11. *ilba*	LX≈PR	FT-BIN	ALIGN-FTO	PARSE	ALIGN-FT
<ilba>	*!			**	
(*) <il>bá		*!		*	σ
☞ (* .) ílba			*		

With these constraints in this order we can correctly derive all Aranda surface patterns. We do not have to adopt an empty vowel in front of consonant initial words and we do not have to make the first segment extrametrical.

We can derive the Alyawarra surface patterns in almost the same way. Remember, however, that Alyawarra was different from Aranda in two respects. Firstly there was the matter of the glide initial words which bear stress on the second syllable in Alyawarra. In section 6.2 we proposed to include Alyawarra initial [w] and [j] in the nucleus, and thus account for the difference with Aranda. A solution based on the same observation applies here. Alyawarra words beginning in [w] and [j] do not start with a consonant. Thus, a foot built over the first syllable violates ALIGN-FTO. Therefore, this first syllable is skipped in the building of feet.

The second difference between Aranda and Alyawarra had to do with the position of stress in disyllabic words that begin with a vowel. Alyawarra stresses these disyllabic words on the second syllable, while in Aranda stress is on the first. From this we may conclude that the ordering of the constraints that are relevant to these disyllabic words is different. In Alyawarra it seems more important to Align the left edge of a foot to an onset than to avoid degenerate feet. This means that the constraints FT-BIN and ALIGN-FTO have to trade places. Let us see what happens if they do.

12. *iylpa*	Lx≈Pr	Align-FtO	Ft-Bin	Parse	Align-Ft
<iylpa>	*!			**	
☞ (*) <iyl>pá			*	*	σ
(* .) íylpa		*!			

It appears that we are able to derive all Alyawarra surface forms with a slightly different ordering of the constraints. Because Align-FtO is now higher ranked than Ft-Bin the foot binarity constraint may be violated, hence the degenerate foot on the second syllable of the selected candidate in (12). This different ordering of the constraints does not affect the analyses of the other words, only in disyllabic words are Ft-Bin and Align-FtO in direct relationship. In longer words Align-FtO is always the only constraint that determines the placement of the first stress. Remember also that the stress placement in Aranda disyllabic words formed the reason to rank Align-FtO under Ft-Bin in the first place. Note also that degenerate feet can only occur in Alyawarra disyllabic words. In larger words Lx≈Pr can be satisfied through the building of a binary foot. Once that is the case degenerate feet are no longer allowed because Ft-Bin dominates Parse.

6.2.3 Conclusion

In this section we have been able to handle the problem of onset sensitivity in Western Aranda and Alyawarra with the assumption of only one new OT constraint. The new constraint, Align-FtO, is crucial in these cases, but it is also of importance to other languages, since it aligns the foot to the onset, and thus prevents the occurrence of “loose” onsets where the foot might have been wrongly aligned to the nucleus. We have shown that, at least in Aranda and Alyawarra, onset sensitivity does not have anything to do with weight, contrary to what was assumed in earlier accounts of the Aranda and Alyawarra data. It rather seems to be a matter of an alignment constraint that interferes with the assignment of stress, where in most other languages it does not. In chapter 5, we introduced many more Australian languages that have stress patterns similar to those of Aranda and Alyawarra (cf. section 5.3.1 and appendix D-VIII). The misalignment analysis that was presented above applies to all of these languages. Thus, no language remains for which we would need moraic onsets, and the weightlessness hypothesis of the onset is saved.

Some other languages from chapter 5 in which the onset seemed to be relevant were Mathimathi, Djapu and Ngarigu. We stated there that the onset-sensitive rules we could adopt to analyse these languages are not possible in our view. Therefore, we need to present alternative metrical analyses for these languages. In the next section such an analysis is presented for Mathimathi. We also refer to Ngarigu, for which the analysis is probably closely related. We offer no analysis for Djapu, but note that it must be possible to derive its stress pattern in a similar vein.

6.3 Mathimathi stress[8]

Mathimathi is an extinct Kulin language that was spoken near the border of New South Wales and Victoria (Australia). Since Mathimathi coronal onsets play a key role in this section, the consonant inventory of the language is presented in table I with the coronals in bold. This inventory is taken from Hercus (1986). (Hercus represents the interdental plosive as d̪, which I replaced with the IPA symbol for its most frequent allophone θ.)

Table I: *Mathimathi consonant phonemes*

	labial	interdental and palatal		alveolar	retroflex	velar
plosives	b	**θ**	(dj)	**d**	**ɖ**	g
nasals	m	**n̪**	(nj)	**n**	**ɳ**	ŋ
laterals		**l̪**		**l**	**ɭ**	
rolled				**r**		
semi-vowels	(w)	**y**				(w)

Only the separation between the boldfaced phonemes and the remaining ones is important for the stress system of Mathimathi. A thorough discussion of the allophones of these phonemes and their distribution can be found in Hercus (1986).

The vowel phonemes that occurred in Mathimathi were /i/, /e/, /u/ and /a/ (diphthongs are /au/, /ai/ and /ui/). These vowels showed no phonemic length distinction. Hercus notes that vowels tended to be phonetically longer under main stress and marginally longer under secondary stress (1986:112). She also mentions sporadic lengthening before nasals, and /g/ and /b/.

---

[8] This section is a rewritten version of Goedemans (1997).

At first sight the Mathimathi stress pattern seems rather straightforward. As we can see in (13a,b), disyllabic words have initial stress. Longer words seem to have initial main stress and a secondary stress on the third syllable when that syllable is non-final or closed (13c,e,f), but not when it is final and open (13d). The data presented in this section are all taken from Hercus (1986).

(13)	a.	bíɳa	‘to go out’
	b.	wáni	‘boomerang’
	c.	gágilàθa	‘to go hunting’
	d.	bérbaθa	‘to jump’
	e.	géŋginìn	‘(your) uncle’
	f.	búndilàθa	‘to go on biting’

A metrical account for this pattern could easily have been devised, were it not that the Mathimathi stress pattern is complicated considerably by the facts in (14).

(14)	a.	dìbárgimàθa	‘to adhere’
		wàuwúnmaθa	‘to be full’
		wùríŋgi	‘grass’
	b.	θàlíŋi	‘tongue, word, speech, language’
		wìθíwaθa	‘to come back’
		dùáŋi	‘glider (flying possum)’
		lìáŋin	‘(your) teeth’
	c.	bérgulù	‘time sticks’
		wìráguθì	‘frill-neck lizard’

d.
```
   x  x     x     x      x
# h h  # h l  # l h  # l l_cor   # l  l
```

The words in (14a) show that Mathimathi stress seems to be quantity-sensitive. The main stress goes to the second syllable if it is heavy (i.e. contains a diphthong and/or a coda consonant), no matter whether the first syllable is heavy or light. Main stress is initial when only the first syllable is heavy, or when the first and the second syllable are light, as in (13c). However, in the latter case there is one regular exception. When both the first and the second syllable are light in the traditional sense (i.e. have a short vowel and no coda consonant), second syllables that have a coronal onset or begin with a vowel in hiatus act as if they are heavy. The words in (14b) have stress on the second syllable though this syllable is clearly not heavy. We also observe in (14) that a secondary stress remains on the first syllable in words that put main stress on the second syllable. (14c) shows that final syllables do not have to be closed to be rhythmically stressed. Finally, the full range of Mathimathi primary stress positions is schematised in (14d), in which the suggested heavy (h)-light (l) distinctions are incorporated.

When normal heavy syllables are involved, stress assignment is fairly straightforward, so we concentrate on main stress in the last two cases in (14d) for the moment. We may try to capture these facts in several ways. Davis (1985) mentions Mathimathi as a case for which onset weight is needed. He claimed that coronal onsets make Mathimathi syllables heavy, just as any coda consonant does, collapsing the third and the fourth case in (14d). In the previous chapters we have seen that such moraic-onset analyses are untenable. Davis realised this, it seems. In a revision of his earlier analysis he captures Mathimathi stress with the help of a *stress shift* rule (Davis 1988). In this analysis coronal onsets can cause a stress shift from the first to the second syllable, accounting for the basic Mathimathi stress patterns.[9] Other approaches to analyse the odd stress assignment rule for the Mathimathi #l l case are available, though.

As was noted in section 5.3.2.3, Goedemans (1993, 1996a) uses *prominence* (see 5.2.3 for an introduction to this phenomenon) to analyse the Mathimathi stress data. Though the details of these analyses differ considerably, the common factor is that coronal onsets make a syllable prominent, and main stress is located on the second syllable if it is prominent. If both syllables are non-prominent, stress falls on the first syllable. When one of the first two syllables is heavy, stress is assigned according to weight, irrespective of prominence considerations (prominence is subordinate to weight).

What these solutions have in common is that they endow onset consonants with the ability to influence stress placement. Remember that we claimed this could only happen if no onset moraicity was involved and if the relevant onset prominence property was prosodically active. The former holds true for both Davis' (1988) and Goedemans' analyses, but they do not observe the latter rule: place features like [coronal] are not prosodic. So, though descriptionally adequate, Davis' (1988) and Goedemans' (1993, 1996a) solutions still feature a type of onset prominence that is illegal in our model. Apart from that, none of the earlier accounts explained adequately why onsetless second syllables should also attract stress.

Gahl (1996) proposes an alternative analysis of Mathimathi stress that

---

[9] Though it does indeed correctly derive the regular Mathimathi stress patterns, Davis' (1988) analysis is subject to some criticism, which can be found in Goedemans (1993), Gahl (1996) and below.

does not feature moraic onsets or onset prominence. Her solution to the Mathimathi problem will be outlined in section 6.3.2.1, after which an alternative will be presented in sections 6.3.3 and 6.3.4.

First, however, we will present some phonetic evidence that confirms the intricacies of the Mathimathi stress pattern, as described by Hercus (1986). The very nature of the stress pattern might have prompted some phonologists to ignore Hercus' description, labelling it as erroneous. Other phonologists only doubt whether the status of the final secondary stress is phonological or phonetic. The acoustic study described in the next section was conducted to check the validity of Hercus' claims, and to find out more about the status of the word-final stress.

6.3.1 The phonetics

Conducting a phonetic experiment on a language that has been extinct for over 20 years, involves some obvious difficulties. Fortunately, the Australian Institute for Aboriginal and Torres Strait Islander Studies (AIATSIS) keeps invaluable tape-recordings of extinct Aboriginal languages, often made with the last speaker(s). The phonetic measurements described below were carried out with several of these tapes as source material.[10]

Phonetic determination of where the stress in a word is actually located, is not a straightforward enterprise. A machine can never do better than corroborate what speakers of the language can hear with their own ears. If there are no speakers available, we can, with some difficulty, measure the acoustic properties of which we know from other languages that they are correlates of stress. One of the problems is that there is no unique phonetic cue for stress, and that the relative importance of the cues may vary across languages. Past research has shown that, for English, pitch and duration are the most reliable perceptual cues for stress, while overall intensity (the result of vocal effort) and vowel quality seem to be of lesser importance (Fry 1955, 1958, 1965; Beckman 1986; van Heuven & Sluijter 1996).

Recent developments in the field have shown that spectral slope (the negative tilt of the spectrum measured at the F1 maximum in the formant tracks of the vowel) is a far better indicator of stress than overall intensity (Sluijter 1995; Sluijter & van Heuven 1996; Sluijter, van

[10] Many thanks to Luise Hercus for graciously allowing me to use her field tapes, and to the AIATSIS for providing me with the copies and a quiet listening room. The tapes used were all recorded between 1965 and 1971 at Point Pearce by Luise Hercus. Her informant was Jack Long, a very old man, who was the last Mathimathi speaker.

Heuven & Pacilly 1997). Sluijter shows that the negative spectral tilt of stressed vowels is less steep than that of unstressed vowels and that this is caused by an increase in physiological effort in the laryngeal system.[11]

I follow van Heuven (1987) and van Heuven & Sluijter (1996) in their assumption that pitch is a correlate of accent (the linguistic device used to place a word or syllable in focus) rather than of stress (a structural, linguistic property). This belief is fed by the fact that the stressed syllable of unaccented words can still be distinguished by a combination of longer duration, greater intensity, and full phonetic quality. Thus, the first syllable of the word *testing* in (15) is stressed, but not accented. The word *say* is in focus and bears an accent (see eg. van Heuven 1987).

(15) I told you to SAY testing, not WRITE it.

The fact that pitch has traditionally been found to be a good cue for stress was due to the limited nature of the experiments reported on. In these experiments, the stressed vowels of which the properties were measured were invariably parts of accented words.

In the acoustic study described below we measure pitch, overall intensity, duration and spectral tilt. For the reasons mentioned earlier we expect duration and spectral tilt to be the most reliable cues for stress (assuming, of course, that Mathimathi employs duration and intensity as stress cues). We have also measured vowel quality, though Hercus writes that vowels in unstressed syllables do not show the tendency to weaken towards schwa (Hercus 1986:114). We need to make sure that Hercus is right in this case because, if the vowels weaken towards schwa, the formant values of the vowels will change. Spectral tilt is directly influenced by changes in formant values. Hence, if the quality of the vowels changes under different stress conditions, we must use a correction factor when we measure the spectral tilt.

6.3.1.1 Materials

As mentioned above, AIATSIS copies of Mathimathi field tapes were used as source material. Three tapes were selected on the basis of recording quality. All the tapes we selected contained careful word-by-word elicitation of Mathimathi speech. From these isolated cases, trisyllabic words were selected that showed various stress degrees of the

[11] This physiological effort leads to a steeper trailing end in the glottal pulse. When the glottal pulse is decomposed in its separate frequency components we find an increase in the high-frequency harmonics, because those are the ones needed to describe steep slopes.

vowel /i/ in the three syllabic positions, reflecting the stress patterns: main stress — no stress — secondary stress, and secondary stress — main stress — no stress. The stress patterns, exemplified in table II, are as they were perceived by Hercus. It was not possible to obtain native speaker judgments on stress position from the last Mathimathi speaker (Hercus p.c.). Hercus notes that primary stress is often accompanied by a slight rise in pitch. This probably means that the words of which she determined the stress patterns were also accented in the sense introduced in the previous section.

Table II: *positions of measured /i/'s by stress. For each cell an example is given, x's indicate structural impossibilities.*

	Main Stress	Secondary Stress	No Stress
1st Syllable	wígaθìn 'dead'	wìn̪áŋu 'who, which'	xxxxxx
2nd Syllable	làɳíŋu 'rib'	xxxxxx	yákilà 'to go round looking'
3rd Syllable	xxxxxx	wúgatì 'take IMP.'	bìyáli 'river red-gum'

The reason why /i/'s were used is rather practical. Only for the vowel /i/ were enough instances found to fill every cell with at least five words. For each cell in table II we looked for five words of reasonably good recording quality. Considerable effort was taken to make sure that the words formed a comparable sample of the stress and length patterns in the Mathimathi language. We included pairs like *bíŋgalì* 'carpet snake' and *mìwúru* 'clever man', of which the first /i/'s that we used in our study were transcribed by Hercus with the same length mark, but which showed a difference in stress level on the first syllable. Another example is the word *wígaθìn* in the first cell of table II. The secondary stressed /i/ in the final syllable would only have been used in our study if this word could have been matched with another word, also ending in /n/ which did not have a secondary stress on the final syllable (which was not the case, so only the first /i/ was used). Such pairs should ensure that statistical biasing effects that might be caused by contextual phonetic lengthening processes are cancelled. The selected words were digitised (22050 Hz, 16 bit) from a REVOX tape recorder and stored on disk.

6.3.1.2 Method

Measurements on the sample files were carried out using the CSL phonetic software package. Durations of the relevant /i/'s were measured on the basis of the visual criteria described in Rietveld & van Heuven (1997). Intensity was measured at its peak in the vowel. Formant values and spectral slope were measured at the same point. Since pitch-contours are not a local property of syllables (which might be a reason for not trying to discriminate between levels of stress on the basis of pitch) we measured by hand the pitch excursion in the word that we judged to be associated with the relevant syllable.

The spectral slope measurements were done in the phonetics laboratory of Leiden University using software developed specifically for that purpose.[12] Following the method described in Sluijter (1995) intensity was measured in four contiguous filter bands: 0-0.5, 0.5-1.0, 1.0-2.0 and 2.0-4.0 kHz. The intensity values were measured relative to the maximum output level of a VAX/VMS AD converter. The only deviation from the method that Sluijter describes is that we did not measure the spectral slope at the F1 maximum in the formant track, the reasons being that we wished to measure all the properties of the vowel at the same point and that the formant tracking mechanisms often had problems finding a clear F1 maximum. We expected the positions of the F1 maximum and the intensity maximum to be highly correlated since the opening of the mouth is maximal at the point of maximal intensity and degree of mouth opening is directly reflected by F1. We therefore measured spectral slope at the point where intensity and formant values were also measured.

6.3.1.3 Results and discussion

We can now proceed to check whether the measured data reflect the stress patterns described in section 6.3. In figure 1 the means for intensity (left) and pitch (right) are given for every syllabic position and stress degree. In the left panel it can clearly be seen that intensity is a poor correlate of stress in Mathimathi, as expected. The slight differences in intensity do not distinguish between stress degree or position. The right-hand panel shows the pitch excursions (in Erbs).[13]

---

[12] The actual measurements were done by Jos Pacilly, who also provided valuable input concerning the spectral tilt cue and pointed out some potential problems in the sample files that might influence the spectral tilt measurements.

[13] Erb scaling is now generally accepted as the best perceptual representation of pitch intervals in intonation languages (cf. Hermes & van Gestel 1991).

These pitch-excursion values do not exactly reflect the patterns that were outlined in section 6.3. If main stress falls on the second syllable there is almost no pitch movement on the first syllable, which should carry secondary stress, while initial main stress is reflected by a more modest pitch-excursion than we found for second syllable main stress. This excursion is followed by a second movement of almost the same size on the third syllable. The latter fact seems indicative of a phonologically relevant final secondary stress, but we cannot be sure. The pitch contours of the Mathimathi words we used were very difficult to interpret, so, though the relatively high pitch excursion on the secondary stressed final syllable is exactly what Hercus predicts, we do not wish to draw any firm conclusions from the pitch data. For now we conclude that these values for pitch are due to the confounding effect that accent has on stress (see above) and continue by looking at more robust correlates of stress like duration and spectral tilt.

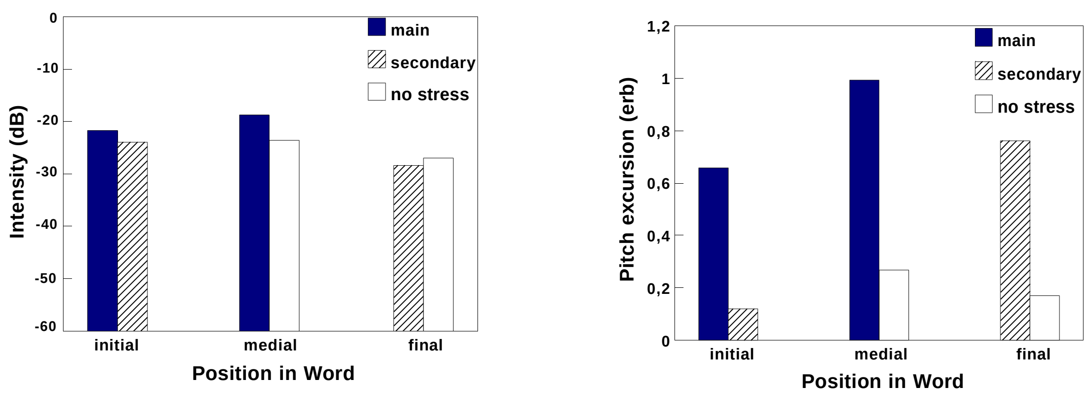


Figure 1: intensity (left) and pitch excursions (right) for the three syllabic positions, broken down by stress degree.

For duration the picture is very clear. As can be seen in figure 2, the mean duration values directly reflect the stress patterns we are looking for. Durations of main stressed /i/'s in both initial and medial position are longer than durations of secondary stressed /i/'s in initial and final position. The latter are longer, however, than unstressed /i/'s in medial and final position.

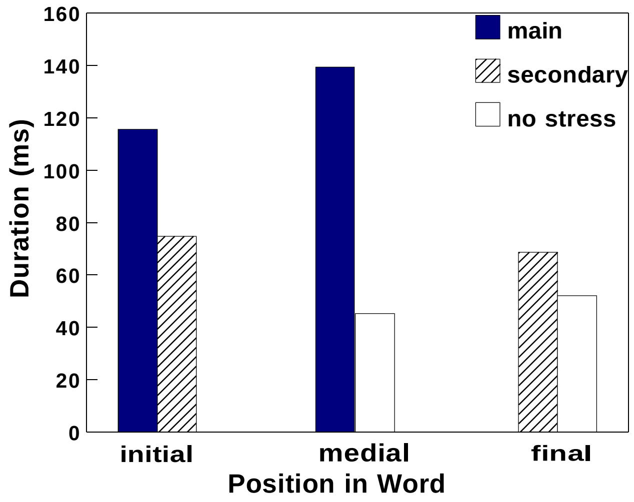


Figure 2: vowel duration in the three syllabic positions broken down by stress degree.

We observe that the means for main stress in initial and medial position are comparable, as is the case for secondary stress in initial and final position and for no stress in medial and final position. Statistical analyses confirm these observations (main stress by position: $F(1,8) = 3.9$, ins., sec. stress by position: $F(1,9) = 0.5$, ins., no stress by position: $F(1,9) = 2.8$, ins.). This allows us to ignore the position factor and use the overall means to show that durations for main stress, secondary stress and no stress are significantly different from each other. A oneway analysis of variance (with SNK post hoc method) shows that they are: $F(2,29) = 62.0$, $p < .001$. These results clearly indicate that there are three degrees of stress in Mathimathi. The highest level (main stress) occurs either on the first or the second syllable, the intermediate level (secondary stress) occurs on the first syllable if the second syllable carries main stress and, significantly, on the final syllable if the first syllable carries main stress. Finally, the third (unstressed) level occurs in medial and final position in the predicted cases. Thus, the duration data confirm Hercus' (1986) claims concerning Mathimathi stress patterns.

The spectral tilt measurements reveal a pattern similar to that found for duration. As noted above, we expect the negative spectral tilt to be steeper for unstressed vowels than for stressed vowels. This means that the spectral curve, which usually has higher values at the lower end of the frequency scale and drops at the higher end of the scale, drops less drastically in case we are dealing with a stressed vowel; there is more energy in the higher regions of the spectrum. In figure 3 we present the spectral tilt results.[14]

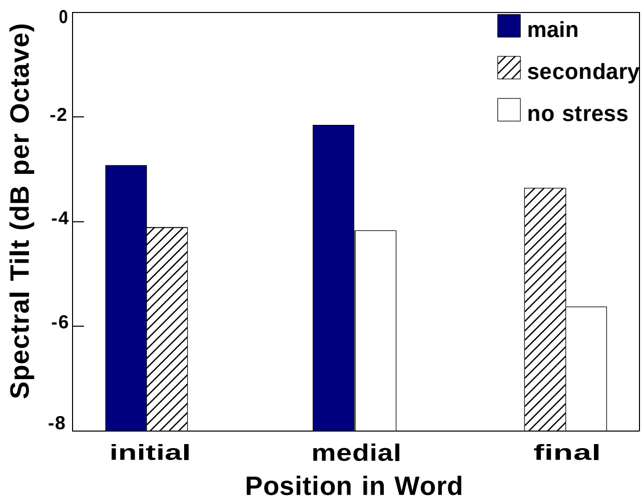


Figure 3: spectral tilt for the three syllabic positions broken down by stress degree.

In this figure we can clearly see that the differences between the three degrees of stress in Mathimathi are directly reflected by the steepness of the spectral curve. Comparison by position shows that the main stress curve is less steep than the curves for secondary stress and no stress in initial and medial position respectively. The difference between secondary and no stress in final position, which is of special interest here, is most clearly reflected by the differences in steepness of their spectral curves.

If we again abstract from the position factor and calculate the overall means for the intensity in the filter bands we find that the intensity values for the different stress levels are close together in the lower end of the spectrum, but significantly different in the upper end of the spectrum. This is confirmed by the statistics: ($F(2,29)=0.5$, ins. [Band 2]) ($F(2,29)=5.6$, $p<.009$ [Band 4]), though a post hoc SNK analysis shows the effect for Band 4 to be mainly caused by the value for main stress

[14] In the method section we stated that there were measurements for four filter-bands. As noted earlier, the measurements for F1, located in the first band, were very unreliable, since they often had quite large bandwidths. This is a characteristic of older speakers. Since the quality of F1 directly influenced spectral slope measurements through a possible interaction with F0, we discarded the first band and calculated spectral tilts from the remaining three bands.

with respect to those for secondary and no stress.

Sluijter (1995) rightly notes that, if one wishes to compare spectral slope measurements of different vowels, these must be corrected for differences in the location of the formants. If the formants of the vowels we wish to compare are located in different positions, this naturally leads to differences in the energy distribution over the filter bands, which might explain variation in spectral slope. What we must do, then, to validate our spectral slope data, is show that the quality of the vowel /i/ is not affected by stress in Mathimathi. If the vowels we used are not significantly different from each other, then the tilt values we found do not have to be corrected for vowel quality. Table III shows the first three formants for the vowel /i/ in each of the three stress conditions.

Table III*: values of the first three formants (in Hz) of /i/ by stress level*

	Main Stress	Secondary Stress	No Stress
F1	348 (60)	364 (60)	379 (85)
F2	2006 (202)	1807 (299)	1837 (284)
F3	2557 (121)	2616 (136)	2577 (192)

A statistical analysis shows that the formant values do not significantly change over the three different stress levels; F1: $F(2,29) = 0.5$, ins., F2: $F(2,29) = 1.6$, ins. F3: $F(2,29) = 0.4$, ins. Thus, as Hercus has already noted, there is no significant reduction to schwa in unstressed vowels. For us that means that no formant-shift correction is needed in the spectral slope data.

Hercus' (1986) claims about the Mathimathi stress patterns are thus confirmed by the phonetic data used in the present study. As expected, differences in intensity and vowel quality are insignificant, pitch is unclear, but differences in duration and spectral tilt reflect the pattern adequately. This leaves us with a phonological problem. We must find an explanation for the odd Mathimathi stress patterns. The rest of this chapter describes an attempt at such an explanation. The next section introduces a line of argument which may provide a solution to the Mathimathi problem.

6.3.2 The morphological component

In section 6.3 it was argued that an onset-insensitive analysis of Mathimathi stress was to be preferred over an onset-sensitive analysis. Gahl (1996) shows convincingly that such an onset-insensitive analysis can readily be found if we recognise the role of morphological structure in the Mathimathi stress rule. She observes that Mathimathi roots are either monosyllabic or disyllabic. In (16) some of the words mentioned in the introduction are repeated with their morphological analysis indicated to illustrate Gahl's claim.

(16)	*verbs*	*nouns*
	bín+a	wán+i
	béb+aθ+a	bérg+ulù
	gág+il+àθ+a	dùá+ŋi
	wìθíw+aθ+a	θàlí+ŋi
	wàuwún+m+aθ+a	wùríŋg+i

(-*a*, -*aθ*-: verbal present tense marker and stem extender, -*il*-: continuative suffix, *(ŋ)ai*, *(ŋ)in* and *(n̪)u*: possessive markers, *(ŋ)i*: intransitive nominative ending, -*m*-: completive marker, -*ulu*-: dual ending)[15]

The generalisation that Gahl (1996) makes is that words with monosyllabic stems have initial stress and words with disyllabic stems have stress on the second syllable, irrespective of whether that syllable is closed (heavy) or starts with a coronal consonant. Thus, her claim is that Mathimathi stress is stem-final, except in disyllabic words (which are invariably stressed on the initial syllable). She proceeds to show that words that would have been exceptions in the onset-sensitive analyses become quite regular, as exemplified by (17).

(17) wàuwúnaθa 'to swell up' wauwun+aθ+a

According to the analyses mentioned in the beginning of section 6.3, this word should have been stressed on the first syllable, since that syllable is heavy and the second is not. The morphological analysis, however, shows that the word has a disyllabic stem, and must therefore be stressed on the second syllable. Hence, only Gahl's analysis derives the correct stress pattern.

An apparent problem for Gahl's analysis comes from the words in (18) which have monosyllabic stems and, nonetheless, have main stress on the second syllable.

[15] There are several other suffixes that play a role in Mathimathi nouns and verbs. These other suffixes follow the general pattern. The suffixes mentioned here suffice to illustrate Gahl's analysis.

(18) bùlg+ái+aθ+a ‘to be soft’
gàθ+íw+aθ+a ‘to flow’

These words are of particular importance since the onset-sensitive analyses do predict the correct stress patterns. The first word in (18) would get main stress on the second syllable because it is heavy and the second word would be stressed such because its second syllable starts with a coronal onset. Gahl’s analysis wrongly predicts main stress to fall on the initial syllable, but she shows that there are compelling reasons to view the words in (18), and others like them, as denominal verbs. The suffixes *-ai-* and *-iw-* serve to derive verbs from the noun stems *bulg+i* ‘soft’ and *gaθ+ini* ‘water’. These derived stems can then take the regular verbal suffixes *-aθ-* and *-a*. The claim then is that these derived stems also take stress on their final syllable. This is the reason why Gahl claims that Mathimathi stress is *stem*-final, and not *root*-final.

In the last section of her paper Gahl demonstrates that the success of the onset-sensitive analyses does not mean that an onset-sensitive stress rule really operates in Mathimathi. She observes a striking asymmetry between the distributions of consonants across positions in Mathimathi. It appears that stem medial consonants in disyllabic stems are almost invariably coronal. Therefore, Gahl claims that the correlation between second-syllable stress and coronal onsets is only apparent: *“it arises because stress is stem-final, and because of the distribution of coronals in stems”* (Gahl 1996:340).

6.3.2.1 The final word?

Since Gahl’s analysis does not respect the weight or prominence of onsets, and accounts for the stress patterns of most Mathimathi words (even the ones that were exceptions in the other analyses), we seem to be allowed to remove Mathimathi from the short list of onset-sensitive stress languages. That is indeed the case, but not before we have noted, and solved, some problems that remain when we adopt the morphological version of the Mathimathi stress rule.

From a typological point of view the claim that Mathimathi stress is stem-final is quite problematic. A survey of the stress systems in Aboriginal languages (appendix D) shows that primary-stress rules referring to the final syllable of the word do probably not occur on the Australian continent. Of course, second syllables may get stress in disyllabic words, or the last vowel of a word may be the only long one in a language that places main stress on the first long vowel. In these cases main stress may be final, but that is accidental: no reference to final syllables is needed to derive these patterns. In the majority of Aboriginal languages a stress rule operates that places the main stress on the initial syllable or near it. Perhaps the penultimate stresses that we find in some “top end” languages of the Northern Territory (geographically far removed from Mathimathi) are as close as the rules can get to the right edge of an Aboriginal word or morpheme. Therefore, an analysis that claims that main stress is on the final syllable of any phonological or morphological unit in an Aboriginal language is suspect.

Furthermore there are words with trisyllabic stems like *bàndálaiθa* ‘to be wet’ which should get stress on the third syllable if the rule is to stress the stem-final one. The fact that this does not happen weakens the claim of stress being stem-final.

A third problem concerns the history of the Mathimathi stress pattern. In an earlier stage, the language must have shared the simple pattern of initial stress with the other Kulin languages. At some stage stress must have shifted to the second syllable of words that had disyllabic stems. But why? The assumption that coronal onsets triggered the shift would undermine Gahl’s theory and put us back to square one. Gahl does not properly address this problem. We will provide a possible reason for this change in the Mathimathi stress pattern in section 6.3.3.

The evidence presented above strongly suggests that Mathimathi stress is not stem-final. At the same time, however, the observations Gahl makes about the correlation between the length of the stems and the position of the stress seem all too valid. In sections 6.3.3 and 6.3.4 an attempt is made to reconcile Gahl’s observations with the idea that stress is not stem-final. In section 6.3.4 a further problem is addressed that is only marginally discussed in Gahl’s paper; that of the secondary stresses. In section 6.3.1 it was shown that the final stress has the same phonetic status as the initial secondary stress, and that these are significantly different from unstressed syllables. The analysis we will present below also accounts for the positions of these secondary stresses.

6.3.3 What do the neighbours say?

To fully understand the solution to the Mathimathi problem presented below, one must know some things about the linguistic area in which the language was spoken. Mathimathi was the northernmost of the Kulin languages, and it formed a sub-group with Ledjiledji (to the west) and Wadiwadi (south). Close to the north of it was Baagandji (Hercus 1982). With respect to the relation of Mathimathi to Baagandji Hercus writes:

"There can be no doubt that Mathimathi belongs to these [Kulin] languages, and yet it shows many interesting features that make it to some extent a transitional language between the Victorian Kulin languages and the language of the Darling River, Baːgandji" (Hercus 1986:101).

The major Kulin languages that were spoken south of the Mathimathi area were Weɽgaia and Wembawemba, and to the east Narinari was located (Hercus 1986). In figure 4 the relevant linguistic area is sketched.

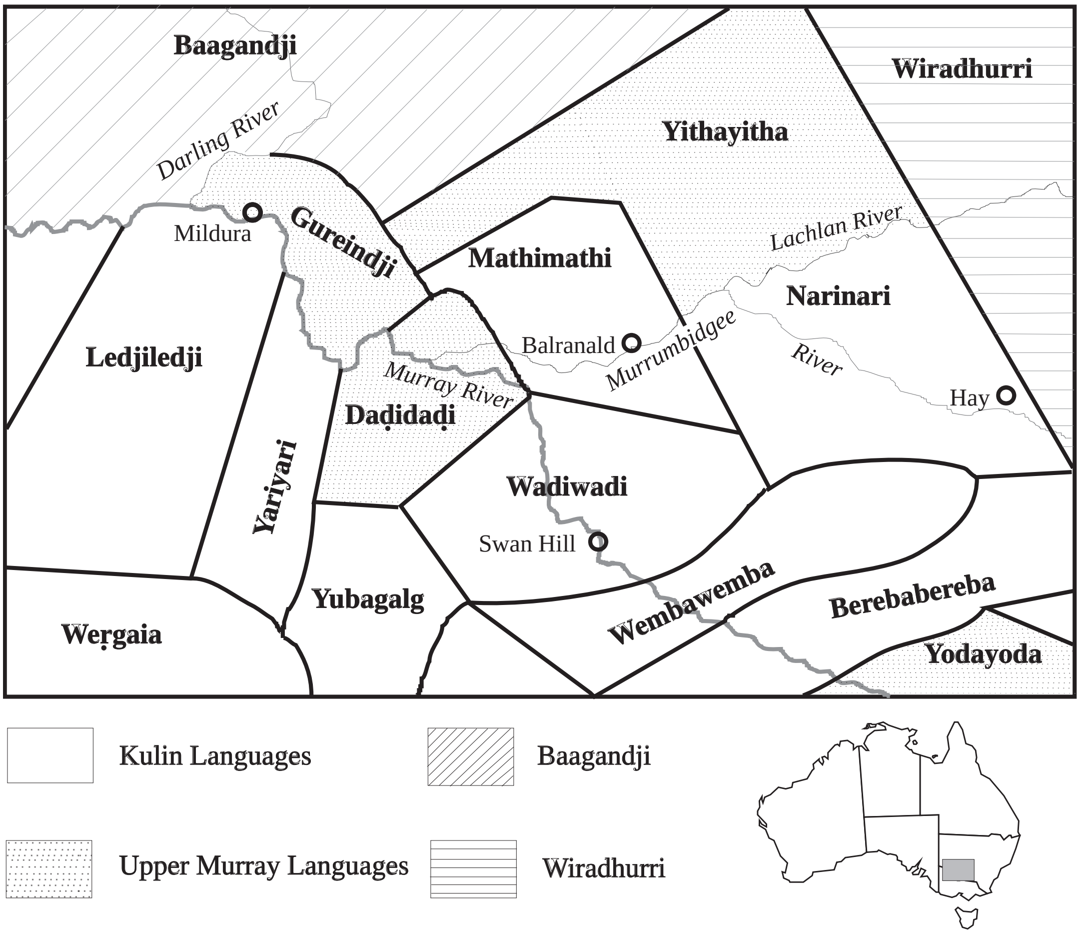


Figure 4: languages in the Mathimathi area (adapted from Tindale 1974; Hercus 1986, 1989)

The two languages that were spoken to the northeast and southwest, Yithayitha and Daɖidaɖi are claimed to be one and the same Upper Murray language (Hercus 1986, 1989; Dixon 1980). As such, the Yithayitha-Daɖidaɖi language is unrelated to either the Kulin group or Baagandji. This language will appear to be of crucial importance to our account of the origin of the Mathimathi stress pattern.

It has often been noted (cf. Dixon 1980) that the rate of borrowing, language split and merger among the Aboriginal tribes was rather high. The linguistic situation was further complicated by the fact that bilingualism was extremely common in the Australian area. These factors have contributed to a situation in which many linguistic features have diffused, sometimes over vast distances across the Australian continent.

With regard to these observations, Mathimathi was a case in point rather than an exception. It was in prime position to have extensive contact with Yithayitha-Daɖidaɖi and Baagandji. The mother of the last Mathimathi speaker, Jack Long, was of part Daɖidaɖi origin, reflecting such contacts. Dixon (1980:258) notes that the languages of the Upper Murray, amidst which Mathimathi is located, presumably had been in contact for a long time and formed a compact diffusional area. Furthermore, Jack Long could "join in" when people spoke Wembawemba and he could understand Yithayitha-Daɖidaɖi. It is in the light of this "fuzzy" linguistic situation that we must see the hypotheses presented below. Let us first see, however, if we can learn something from a comparison of Mathimathi with the two major Kulin languages.

6.3.3.1 Nominative /i/

A close comparison between Mathimathi, Wembawemba and the Djadjala dialect of Weɽgaia (Hercus 1986) provides a list of cognate nouns, a selection from which can be found below.

(19)	*English*	*Mathimathi*	*Wembawemba*	*Weɽgaia (Djadjala)*
a.	brolga	gùθúni	gúθun	gúdjun
	white cockatoo	djìnáwi	djínab	djínab
	crow	wáŋi	wá	wá
	boomerang	wáni	wán	--
	father	mámi	mám	mám
	fire	wànábi	wánab	wánjab
	sand	gùrági	gúreg	gúrag
b.	teal duck	béɳer	béner	béɳeɽ
	canoe	yúŋwib	yúŋwidj	yúŋwib
	cod	bán̪d̪el	bándjil	bándjil
	poison	θándel	θándel	--
	reed mace	gámbaŋ	--	gámbaŋ
	rush	búŋeɖ	búŋuɖ	búŋuɖ
	spear pointed waddy	bérbin̪	bíɽbenj	bíɽbinj

Many Mathimathi nouns have the obligatory intransitive nominative ending *-(ŋ)i* when pronounced in isolation. It appears that Wembawemba and Djadjala nouns do not share this feature. By far the majority of Mathimathi words act like the words in (19a). There are two possible reasons for this behaviour:

1. The *-(ŋ)i* was added to monosyllabic nouns to eliminate all monosyllables from the language, possibly under the influence of Baagandji which had no monosyllables.
2. The *-(ŋ)i* was added to avoid word-final stress (resulting from a shift of stress to the second syllable of disyllabic words), a well documented tendency of the world's languages (cf. McCarthy & Prince 1993).

Later, addition of *-(ŋ)i* might have been extended to other nouns.

Both Wembawemba and Djadjala also have an ending - *i*, but it is reserved for the vocative, for which it is also used in Mathimathi.[16] It is possible that the endings are related and that its usage was extended beyond vocative to nominative in Mathimathi.

It may be noted that verbs do not follow the pattern outlined in (19). The Wembawemba and Djadjala forms are generally the same as those in Mathimathi (apart from frequent addition of the stem extender *-aθ-* in the latter). This is due to the obligatory present tense marker *-a*, which makes sure that monosyllabic verbs do not exist and that stress is never word-final in verbs.

The words in (19b), which represent a minority, attract our attention, since they clearly have disyllabic stems. According to the analysis outlined in section 6.3.2 they should, therefore, have the main stress on the second syllable. One could object that this main stress would then be word-final, but a strategy to avoid this word-final stress is readily available. In fact, the disyllabic stems in (19a) have followed this strategy and added the ending *-i* to avoid word-final stress. Why some Mathimathi disyllabic stems have added *-i* to enable a shift of stress to the second syllable without violation of the word-final stress constraint, and some have not, remains a mystery under an analysis that is as rigid as Gahl's stem-final stress proposal. This is considered to be another drawback of her analysis. The proposal presented below contains an explanation for these facts.

[16] In Wembawemba the vocative was only marked by the -i suffix in kinship terms. An interesting detail is that the normal vocative marker was a strong secondary stress on the final syllable, eg. *wílkàr* 'hey, dingo' (Hercus 1986:29).

6.3.3.2 Stress patterns

The languages surrounding Mathimathi all have the “initial main stress” pattern that is so common to the Australian area (though Baagandji has some odd characteristics that makes it stand out from the rest, Hercus p.c.). Yet, some of these languages also have words with exceptional stress patterns. These patterns are very reminiscent of the Mathimathi pattern, as can be seen in (20). Data are taken from Hercus (1986).

(20)	*Wadiwadi**		*Narinari**	
	gàráwi	‘big’	θùgúli	‘whitefellow’
	mìlági	‘dust’	gàyíni	‘water’

Wembawemba

brídjìrim	‘resin from the Murray pine’
gánjèŋga	‘to cough’
mádèmbola	‘to call as a witness in a trial’
wírèŋgal	‘fish, the callop, yellow-belly’

Of the languages marked with * very little remains, so we can never be sure of the extent to which this pattern was present. Yet, the general picture seems clear. Most of the languages in the Mathimathi area have words in which main stress can fall on the second syllable, but none of them seems to employ this feature as much as Mathimathi. Certainly none of the other languages would have a stress *rule* stating that main stress goes to the second syllable under certain conditions. There are simply too few words that have this pattern.

As noted above, Yithayitha is of special importance. Unfortunately, documentation on Yithayitha and the other Upper Murray neighbours of Mathimathi is extremely scarce. However, both Dixon (1980:46) and Hercus (1986:101, 1989:59) mention Yithayitha as the closest relative of Yaralde (McDonald 1977), spoken at the mouth of the Murray, which is one of the few well documented Murray languages. A legend about the migration of the Yaralde tribe from the Upper Murray to the Murray mouth suggests that Yaralde might once have been spoken in the same linguistic area as Mathimathi. It has been suggested that the Murray languages formed a dialectal chain all the way from Yaralde at the Murray mouth to Gaanay/Kurnai in Gippsland (Hercus p.c.).[17]

---

[17] Taplin (1879) notes that, along the Murray, members of every tribe knew the languages that were spoken directly upstream and downstream, so that messages could be sent from the Murray mouth all the way upstream.

The nature of the Yaralde stress rule, as described by McDonald (1977), sheds new light on the Mathimathi problem. As in all the languages discussed so far the default position for main stress in Yaralde is the initial syllable, with secondary stresses occurring on alternate syllables thereafter. In some cases, however, stress shifts over a single apical sonorant, leaving disyllabic monomorphemic roots in which stress is equally distributed over the first two syllables. In (21) we present some examples.

(21)	a.	wrép	‘grind’	tɭáyp	‘spread’
		mɳákur	‘reed’	kríŋunàl	‘plant’
	b.	kálát̲	‘dawn’	ŋúnáp	‘lizard’
		ŋáwánt̪	‘camp’	tánjákaw	‘axe’
		ŋánán	‘pigface’	párráyi	‘honey’
		ŋáyámp	‘back’	táréym	‘cut around’

This phenomenon is related to initial syllable loss. According to McDonald the stress shift facilitates deletion of the first vowel, and thus, formation of initial clusters, as in (21a). The process seems to have been partially blocked by resisting initial /a/’s and a restriction on the formation of impossible initial clusters. These blockings account for the words in (21b).

We propose that the Mathimathi stress rule has arisen through indirect contact with Yaralde, through neighbouring Upper Murray languages like Yithayitha. In the Victorian linguistic melting pot, as it was described in section 6.3.3, this would not have been unlikely. A possible course of action might have been that Mathimathi borrowed the Yaralde type stress shifts, later phonologising the exceptional pattern to a rule that put stress on second syllables if that was morphologically possible. Hence, Mathimathi adopted the Yaralde stress pattern in a slightly modified form: stress shifted in disyllabic stems, but the shift was disallowed when it would place primary stress after a (monosyllabic) stem.[18] We will outline the details of our analysis in section 6.3.4. Let us first look at some arguments in favour of the view that the Mathimathi stress pattern changed under the influence of its Upper Murray neighbours.

One reason why we believe that the Mathimathi pattern arose through diffusion is that it explains why the pattern is not applied across the board. Even in Gahl’s account there are many unexplained exceptions. In

[18] It may either be the case that Hercus’ description of Mathimathi captured the language in a transitional phase and that it was going the same way as Yaralde, or that Mathimathi borrowed the stress shift rule only, deletion of initial vowels never being the objective.

a language contact situation exceptions are expected to occur. This becomes even more apparent in Wembawemba, Wadiwadi and Narinari, which only marginally show reflexes of this unusual stress pattern (whether the languages borrowed from Yaralde in their history or from Mathimathi in a later stage remains an open question).

Other evidence comes from Mathimathi itself. Some exceptions have stress patterns that more closely reflect the Yaralde situation, or suggest that the relation between primary stress on the second syllable and secondary on the first is not as rigid as it would seem. Some of these words are given in (22).

(22)	búyíŋga	'to blow (a fire)'	dágθèraθa	'to fight'
	díríli	'sky, heaven'	mún̪ábi	'round knob waddy'

Furthermore, we observe that Yaralde and the Kulin languages have more traits in common. An example of this is the ending -i, discussed in section 6.3.3.1. Yaralde employs this ending to mark the nominative and accusative forms of the singular pronouns, and optionally adds it to nouns in their singular form.

More evidence comes from two languages even further up the Murray river; Yodayoda and Ngarigu (Hercus 1986). Of these two, Yodayoda is related to Yaralde, and Ngarigu is spoken to the east of Yodayoda. Yodayoda shows exceptions that are similar to those we find in Wembawemba and, as we have seen in section 5.3.2.3, main stress in Ngarigu falls on the second syllable if that syllable begins with /d, r, l, m or n/. The examples in (23) illustrate the patterns.

(23)	*Yodayoda*		*Ngarigu*	
	málòga	'sand'	guníring	'useless, silly'
	yágòrumdjak	'come here'	djarímiŋ	'happy, flash'
	gúθùbna	place name	budálag	'tree-goanna'

We can see that Ngarigu closely resembles Mathimathi, though few examples can be found. We consider the fact that Ngarigu shows exceptional stress patterns similar to those in Mathimathi and its neighbouring Kulin languages, while it also neighbours an Upper Murray language, to be suggestive of a similar borrowing situation.

Evidence of a less well established nature comes from Ngayawang, an Upper Murray language located between Yaralde and Yithayitha. Moorhouse (1846) does not directly address stress patterns in Ngayawang, but notes that vowels bearing the *length accent* are marked with "-", as in (24).

(24)	parētun	'to grease'	pimēllinki	'flowers of the wattle'
	tortōko	'selfish, avaricious'	tummūn	'to seize'

In Moorhouse's word list, these length accents only occur on the first or the second vowel of the word. We cannot be sure that they represent stress, but they are at least indicative of the importance of the first *and* the second syllable in Ngayawang. This indicates that placing stress on the second syllable may indeed be an areal feature of the Murray languages.

Finally, though Djadjala is closely related to Wembawemba, as evidenced by many common lexical items such as *wireŋgal* 'fish', Djadjala does not have a single exception to the initial main stress rule. Again, a diffusion account explains this fact, since Djadjala is further removed from the rest, and thus less likely to be influenced by the Upper Murray languages or Mathimathi.[19]

If we accept the claim that the Mathimathi stress pattern arose through contact, all the problems that we had concerning the analysis presented in section 6.3.2 vanish. It is no longer the case that stress is stem-final (though morphology is still involved, as we will see below). Hence, the objection that finality involved in a stress rule of an Aboriginal language is suspect is no longer applicable. Stress moving to the second syllable for various reasons is a well attested feature of Aboriginal languages (see sections 5.3.1 and 6.2). Similarly, we no longer have to explain why trisyllabic stem words have no stem-final stress. The stress can shift to the second syllable, but no further.

The borrowing solution also explicitly explains how the Mathimathi pattern arose. One could object that an onset-sensitive rule would then still be needed to explain the historical developments in Yaralde. It is very likely though, that the stress shift in Yaralde was driven by motivations other than the identity of the onset of the second syllable. A strong tendency to reduce the first syllable could have been the trigger for a stress shift.

The Mathimathi case is slightly different. As noted above, we believe that Mathimathi *borrowed* the Yaralde-type stress-shift rule through the

---

[19] Further pressure for Mathimathi to put stress on the second syllable may have come from Baagandji. In this language stress is mainly initial, but, as Hercus (1982) notes, another pattern that cuts right across the basic pattern is the following: in some words consisting of two morphemes of which the second begins with a vowel, a strong accent falls on the juncture syllable. Many of these rising accents fall on the second syllable:

(i)	idja + ulu	[itjôlu]	one + SG	'all alone'
	iinga + aba	[iingâpa]	sit + 1sg sub	'I sit'

Upper Murray languages and later altered it from a shift rule to a rule that placed stress on second syllables wherever the length of the stem allowed that. All this would mean that Davis (1988) was right after all in assuming a stress shift rule, though the motivation for it is completely different. It seems that the shift was not triggered by coronal onsets, but by initial syllable reduction in Yaralde and by diffusion in Mathimathi. The fact remains, as noted at the end of section 6.3.2, that *all* stem-medial consonants are coronal, which allowed Davis to make the generalisation that coronal onsets attract stress. The question why this pattern arose is not easy to answer, but a clue can be found through the comparison with Yaralde presented in this section. In Yaralde, as in many other languages, coronals were the best candidates for the second position in a newly formed onset cluster. Thus, stress shift to the second syllable and reduction of the first syllable was easiest when the second syllable had a coronal onset. It is possible that this was reflected in Mathimathi by a lenition process that reduced all non-coronal stem medial consonants to coronals (or deleted them). As it is, this clue is rather tentative. More (comparative) research is needed to account for the uneven distribution of coronals in Mathimathi stems.

Finally, we no longer have to worry about the facts mentioned in section 6.3.3.1. In a borrowing situation it is very common that there are exceptions to the general pattern. The words in (19b) must have rejected a Yaralde-type stress shift for reasons that remain unclear. This may account for their exceptional position in Mathimathi. After resisting the stress shift (thus not adopting a final -i), they would also reject the general rule placing stress on the second syllable, thus avoiding word-final stress (see 6.3.3.1).

As one may have noted, the matter of the secondary stresses has not been addressed yet. An account of secondary stresses in Mathimathi is integrated in the analysis presented in the next section.

6.3.4 Reanalysis

In the following account of Mathimathi stress I will assume that Gahl's (1996) claim that stress is morphologically determined is basically right. Such a morphological analysis neatly accounts for most of the Mathimathi data. I do not assume, however, that stress in Mathimathi is stem-final. I propose that in adapting the stress rule that we have illustrated through Yaralde, Mathimathi took the morphological consideration of stem length into account. It placed stress on the second syllable only if the stem was long enough to do so. Thus, Mathimathi stress is morphologically determined, but, as in most languages that have variable stress positions, the assignment of main stress is confined to a two syllable window at one of the word edges, in this case the left one. Secondary stress seems to be assigned without such morphological restrictions.

The analysis presented below is again couched in Optimality Theory.[20] We believe that this theory, based on the ranking of violable constraints, is especially well suited to handle the Mathimathi stress rules which are so obviously hampered by dominant morphological and phonological restrictions. Let us, then, look at Mathimathi one final time.

6.3.4.1 A morphologically bound analysis of Mathimathi stress

An OT analysis of Mathimathi must in principle use the same constraints in more or less the same ranking order as an analysis for Aranda and Alyawarra. This is so because the basic stress pattern is the same: main stress on initial syllables and secondary stress on alternates thereafter. Remember from section 5.2.1 that we introduced the undominated constraint FOOTFORM(Trochaic) to ensure the trochaic parsing needed to derive this basic pattern. We assume that this constraint is undominated for Mathimathi as well. We further observe that we will probably need degenerate feet to derive the first secondary stress of some examples in (14), which can never be part of disyllabic feet. Hence, it must be more important to parse syllables than to avoid non-binary feet. This is expressed in the ranking of PARSE above FT-BIN (the reverse of the ranking for Aranda and Alyawarra). The final constraint we introduced in 5.2.1 is ALIGN-FT which we argued must be ranked below PARSE. Otherwise the string of syllables after the first foot would remain unparsed. ALIGN-FT must also be ranked below FT-BIN; if it is more important to align each foot to the left edge of the word than to make binary feet, all feet would be degenerate. Such monosyllabic feet would allow the left edge of the following foot to be one syllable closer to the left edge of the word than disyllabic feet, leading to a less severe violation of Align-Ft. Thus, we obtain the following ranking:

FOOTFORM(Trochaic) >> PARSE >> FT-BIN >> Align(Ft,L,PrWd,L).

We can use this ranking to derive the correct pattern for Mathimathi words with the basic stress pattern, as in (25) (undominated FOOTFORM is not represented).

---

[20] For an analysis in a different framework, see Goedemans (1997). From here on this chapter deviates strongly from that article.

25. *bergulu*	PARSE	FT-BIN	ALIGN-FT
(* .) bérgulu	*!		
☞ (* .)(*) bérgulù		*	σσ

We have deliberately left out a third logical candidate here. Note that all candidates that leave one syllable unparsed will lose immediately. However, a candidate that parses the first syllable in a degenerate foot and the second and third in a binary foot will be more optimal than the selected output in (25) since the second foot violates ALIGN-FT only once. We must eliminate this candidate with a constraint that is ranked above ALIGN-FT. Note that the candidate we propose here is indeed less optimal than the others in one respect. A degenerate foot followed by a trochee means two stresses are adjacent. The standard constraint against such adjacent stresses is NOCLASH. We propose to rank it above ALIGN-FT and leave its ranking with respect to FT-BIN open, hence the dashed line between them. In (26) we eliminate the candidate in question through violation of NOCLASH.

26. *geŋginin*	PARSE	FT-BIN	NOCLASH	ALIGN-FT
(*)(*)(*) géŋgìnìn		**!*	**	σ/σσ
(*)(* .) géŋgìnin		*	*!	σ
☞(* .)(*) géŋginìn		*		σσ

One may have noted that we have deliberately chosen those words here that do have a final stress. From the host of Mathimathi words that do not have a final stress where we would derive it, we could have deduced a prohibition of degenerate feet, ranking FT-BIN above PARSE. However, we already noted that we do need these degenerate feet. So, another solution must be found for the differences in final stressing. We have stated in section 6.3 that this seems to involve quantity-sensitivity. Such QS stress assignment would explain why *géŋginìn* gets a final secondary stress and why *bérbaθa* does not. It does not, however, explain why *bérgulù* gets a secondary stress. In section 6.3.1 we have found evidence for the phonological reality of the final stress, which means that we cannot attribute these effects to phonetics, but that we must incorporate a phonological constraint for final secondary stresses into our analysis.

It seems to us that we can make a generalisation to solve this crisis. A thorough inspection of Hercus' word list reveals that stressing of final light syllables does not occur in verbs or adverbs, and that more than two thirds of the nouns and adjectives do stress these syllables. Moreover, of the 103 unstressed final syllables in verbs and adverbs 86 were -*θa*, 5 were -*ta*, and 6 were -*la* (the rest was idiosyncratic but ended in -*a*). It seems that the class of offending unstressed final syllables all contain the present tense marker -*a*. We propose to employ *extrametricality* to place the present tense marker outside the metrical analysis, but there are other possibilities (like a morphological constraint that places -a outside the PrWd). The relevant constraint would be:

• EM(-a): do not stress the present tense marker -*a*[21]

This constraint must be ranked above parse, so that violation of EM(-a) is worse than having an unstressed final syllable. Hence the evaluation tableau for *bérbaθa* would look like (25) with EM(-a) ranked before PARSE. This will result in the rejection of the candidate with the final stress, before PARSE gets the chance to select it, as in (25). In the tableaus below, PARSE and EM(-a) will play no role of significance. Hence, we will assume they are ranked as proposed, but since they are otherwise undominated we will not represent them. The candidates we will consider below are, therefore, always fully parsed excepting final -*a*, so as to satisfy these constraints.

We can now start working on the patterns that make Mathimathi deviant from a regular trochaic left-to-right system. Since we must obviously distinguish between main and secondary stresses when we look at the more complicated cases, we introduce the constraint that places the head in the prosodic word first. McCarthy & Prince (1993) propose the constraint ALIGN-HEAD which in our case is:

• ALIGN(PrWd,L,H(PrWd),L)

So, the left edge of the prosodic word must match the left edge of the head foot. This constraint can be violated gradually, like ALIGN-FT. Had

[21] This is a language specific constraint. Whether such constraints are possible or not is a matter still under debate. We assume they must exist, but we will not enter the debate here.

we indicated heads in (25) and (26) none of the candidates would have violated ALIGN-HEAD (assuming the heads would be on the first foot). Some of the heads in (14), however, can under no circumstance be aligned to the left edge of the prosodic word. Some other, higher ranked, constraint must force these feet to misalign and violate ALIGN-HEAD. From the discussion above it will be clear that our proposal for this constraint will be the OT translation of the borrowed shift rule. A constraint that nicely expresses the forced shift off (for some reason) unfit initial stress bearers is the version of *NonInitiality* (NONINI) that Visch (1996) uses in her OT analysis of the South American language Carib. It is formulated as:

• NONINITIALITY: the head of a prosodic word is not the leftmost foot.[22]

Both ALIGN-HEAD and NONINI must be ranked above FT-BIN. If ALIGN-HEAD did not dominate FT-BIN, heads of Mathimathi words could never occur one syllable away from the left edge of the word. That syllable would have to be parsed into a degenerate foot which the higher ranked FT-BIN would prohibit. So, we rank ALIGN-HEAD above FT-BIN and NONINI above ALIGN-HEAD (since all heads would be perfectly aligned to the left edge if we allowed ALIGN-HEAD to dominate NONINI). The ranking is thus:

NONINI >> ALIGN-HEAD >> FT-BIN >> NOCLASH >> ALIGN-FT

Below we will no longer indicate ALIGN-FT since it is of no consequence to the analyses presented there. Let us put this constraint ranking to work in the evaluation of some Mathimathi words. The words we are interested in have main stress on the second syllable. We should see NONINI pushing main stress one syllable to the right (no more) violating ALIGN-HEAD minimally. In (27) and (28) we find example evaluations for three and four syllable words respectively. (To save space in the tables the head is indicated by underlining the asterisk in question.)

---

[22] Notice the difference in the level of prohibition between this constraint and Prince & Smolensky's (1993) NonFinality constraint which prohibits *every* form of stress on final syllables while NonIni only disallows initial main stress. We may further note then, that OT is well suited to separate the constraints for the assignment of main and secondary stress (also at a more basic level, cf. Align-FT and Align-Head) which, as defended by Van der Hulst (1984, 1996a, 1996b), is the course metrical phonology must take if we wish to successfully analyse all possible stress systems.

27. *wuriŋgi*	NONINI	ALIGN-HEAD	FT-BIN	NOCLASH
(*)(*̲)(*) wùríŋgì		σ	**!*	**
(*̲)(* .) wúrìŋgi	*!		*	*
(* .)(*̲) wùriŋgí		σσ!	*	
☞ (*)(*̲ .) wùríŋgi		σ	*	*

28. *wiraguθi*	NONINI	ALIGN-HEAD	FT-BIN	NOCLASH
☞ (*)(*̲ .)(*) wìráguθì		σ	**	*
(*̲)(* .)(*) wíràguθì	*!		**	*
(* .)(*̲ .) wìragúθi		σσ!		
(*)(*̲)(* .) wìrágùθi		σ	**	**!

The second candidate in (27) is eliminated immediately because it violates the highest ranked constraint (NONINI). The third candidate in (27) violates ALIGN-HEAD to a worse degree than the other two remaining candidates, so it loses out. Note that we can still not decide on the ranking between FT-BIN and NOCLASH. It is unimportant by which constraint we select the optimal output. In (27) we have chosen to act as if FT-BIN dominates NOCLASH and shaded the irrelevant boxes accordingly for reasons of clarity, but it could just as well have been the other way around. We can eliminate the first candidate either because it violates FT-BIN two times more than the fourth candidate, or because it constitutes a worse violation of NOCLASH.

In (28), candidates of which the head is on the first foot, or those that place it more than one syllable away from the left edge are eliminated in the same fashion as in (27). In this case though, the decision between the remaining candidates cannot be made on the basis of FT-BIN violations. They both violate this constraint twice. The optimal output is selected on the basis of NOCLASH.

With this ranking we have introduced a major problem into the analysis of Mathimathi stress. Now we derive *all* the words with main stress on the second syllable, also the ones in (25) and (26). To prevent this from happening we must translate Gahl's (1996) morphological observation into a constraint. We use McCarthy & Prince's LX≈PR(*Mcat*) and fill the *Mcat* slot with the morphological category *stem*. Thus we get the following extension of the LX≈PR constraint:

• LX≈PR(*stem*): let a morphological stem correspond to a prosodic word.

Prince & Smolensky view the PrWd as "the domain in which 'main stress' is defined" (Prince & Smolensky 1993:43). Hence, we may interpret our LX≈PR(*stem*) constraint liberally as: the head of the prosodic word must fall on the stem. If we do so, and we rank this constraint above all other relevant constraints, our problems are solved. Remember that words with initial stress have monosyllabic stems and words with main stress on the second syllable have disyllabic stems. If we try to derive main stress on the second syllable of a word with a monosyllabic stem we violate LX≈PR(*stem*). Consider (29) in which the stem is embraced by square brackets.

29. *wigaθin*	LX≈PR	NONINI	ALIGN-HEAD	FT-BIN	NOCLASH
(*̲)(* .) [wíg]àθin		*		*	*!
(*)(*̲ .) [wìg]áθin	*!		σ	*	*
☞ (*̲ .)(*) [wíg]aθìn		*		*	

So, words with monosyllabic stems, like those in (25) and (26), get initial main stress by virtue of LX≈PR(*stem*). In words with disyllabic stems main stress may be placed on the second syllable to satisfy NONINI, because this can be done without violating LX≈PR. Hence, the evaluations in (27) and (28) do not change drastically. Note however, that candidates placing the head two or more syllables away from the left edge are now rejected on the basis of a LX≈PR violation and no longer on their worse violation of ALIGN-HEAD. Yet, that does not mean that ALIGN-HEAD has become completely irrelevant. Remember that in 3.1 we noted that *bàndálaiθa* would get main stress on the third syllable in Gahl's analysis, since it has a trisyllabic stem. In our analysis, this trisyllabic stem allows the head to appear on the third syllable without incurring a violation of LX≈PR. The reason why the head is not located on the third syllable in the output is that ALIGN-HEAD is active in keeping the head as far to the left edge of the word as possible (while avoiding NONINI violations, see tableaus 27 and 28).

We can see ALIGN-HEAD at work again in the solution to the final problem concerning Mathimathi stress. We have not yet dealt with disyllabic words. Though an analysis of disyllabic words, which all have initial stress, should be unproblematic, there is a slight complication. Normal disyllabic words like *wáni* should get main stress on the initial syllable because they have a monosyllabic stem. Candidates with main stress on the second syllable are ruled out immediately by LX≈PR(*stem*), as in (30). The subset of disyllabic words in (19b), however, has disyllabic stems. For this subset, LX≈PR does not eliminate candidates that put stress on the second syllable. Hence, NONINI would reject the candidate with initial main stress, which we want as the proper output. This problem can be solved by switching the ranking of NONINI and ALIGN-HEAD, as in (31). The fact that we must rerank NONINI to a lower position for this set of exceptions reflects what we have said earlier about these (19b) words. For some reason they rejected the stress shift rule that is embodied in NONINI.

30. *wani*	LX≈PR	NONINI	ALIGN-HEAD	FT-BIN	NOCLASH
(<u>*</u>)(*) [wán]ì		*		*!*	*
(*)(<u>*</u>) [wàn]í	*!		σ	**	*
☞ (<u>*</u> .) [wán]i		*			

31. *beɳer*	LX≈PR	ALIGN-HEAD	NONINI	FT-BIN	NOCLASH
☞ (*̲ .) [béɳer]			*		
(*)(*̲) [bèɳér]		σ!		**	*
(*̲)(*) [béɳèr]			*	*!*	*

This treatment of disyllabic words brings us at the end of our discussion of Mathimathi stress. Our OT analysis covers all cases while only one constraint was needed that was not previously proposed to be part of the universal set. This was EM(-a) which we, for obvious reasons, cannot possibly do without.

With this constraint based approach we have eliminated quantity-sensitivity from the Mathimathi stress rule altogether. The stem-final rule that Gahl proposed for main stress was already QI, like our analysis. Now it appears that we can also assign secondary stress in a quantity-insensitive fashion. We can simply assign stresses to all the syllables that are in stressable positions, except extrametrical -*a*.

6.3.5 Conclusion

In this final section we have presented phonetic evidence that supports the stress patterns for Mathimathi as they were given by Hercus (1986). This evidence included support for the phonological reality of the final secondary stress. We have further reviewed a proposal by Gahl (1996) in which it is claimed that Mathimathi stress is morphologically determined. The advantage of this proposal is that it eliminates the role of onsets in the Mathimathi stress rules, which is, as we have noted, a gain of considerable theoretical importance. Before we could shorten the list of onset sensitive cases by one, there were some problems with Gahl's analysis that had to be solved. In the final sections of this chapter we have introduced the possibility that the Mathimathi stress pattern might have arisen through contact with its neighbouring languages, and given ample evidence to support this view. Such an explanation for the odd Mathimathi stress pattern, combined with Gahl's observations about its morphological component, removes all the problems we might have with Mathimathi stress. In the last section we showed that the Mathimathi stress rule can be reduced to a relatively simple quantity-insensitive one. We do not present it as the ultimate truth. It may be open to improvements. However, we strongly believe that alternative analyses will generally proceed along the same lines.

We hope that this section, and Gahl (1996), show that whatever the Mathimathi stress rule may be, it is not onset-sensitive. The removal of Mathimathi from the list of onset-sensitive languages brings us one step closer to exceptionless support for the claim that, if onsets can contribute to prominence in the first place[23], only prosodically active onset properties can do so. The feature [coronal] that was supposed to influence Mathimathi stress did not fit this description. We have shown it to be irrelevant to stress, which would be assigned to second syllables of disyllabic stems no matter what the initial consonant of those syllables would have been. It is the result of NONINI pushing stress away from the first syllable if possible. That also explains why onsetless second syllables of disyllabic stems receive main stress. Remember that this was a problem for the earlier analyses.

We have also shown that some exceptional cases of initially stressed words with disyllabic stems can be dealt with by a simple switching of NONINI and ALIGN-HEAD in the constraint ranking, reflecting the relatively low importance that NONINI has in this set. We believe that this move also holds the solution to the Ngarigu stress pattern, in which main stress on the second syllable is exceptional. Such behaviour may very well be expressed in the ranking of NONINI above ALIGN-HEAD for the Ngarigu exceptions while NONINI is ranked below ALIGN-HEAD for the regular cases (or, perhaps, the two contraints remain unranked, which is another way to account for the variation). As such, the ranking of NONINI reflects the shifting of stress from the initial to the second syllable that seems to be an areal feature of the languages spoken along the Murray River.

---

[23] Cf. section 5.4, in which we note the caution with which we must draw conclusions about onset-prominence from the only two remaining cases: Pirahã and Puluwat.

7 Conclusions

7.1 Introduction

This book has been mainly concerned with the role of onsets in stress rules. It has been made clear from the start that this role is virtually non-existent. However, the recognition of this fact is not what constitutes the contribution of this book to the field of metrical phonology. On the contrary, it has since long been known that only the nucleus and coda can contribute to syllabic weight, and thus influence the location of stress in languages that have quantity (weight)-sensitive stress rules. It was observed that only the length of the vowel and/or the presence or absence of coda consonants seemed to be of influence to syllable weight, the onset being inactive, or “weightless”. The absence of a role for the onset in such quantity-sensitive stress rules has never been the issue of much debate, since the rule that syllable weight is independent of the presence or absence of the onset seems to be valid cross-linguistically. Hence, the weightlessness of the onset is considered to be a linguistic universal. In one respect, however, disregarding the onset in the quantity debate is scientifically naïve. In our view, the question *why* the onset cannot contribute to syllable weight remains unasked all too often. Surely some interesting auditory (or psycholinguistic) phenomenon must underlie such a strongly supported linguistic universal. It is the search for this phenomenon that we have undertaken in this book. The central research question that formed the basis for the experiments we conducted was: *Why is the syllable onset weightless?* Our major conclusions regarding this question are summarised in section 7.2.1 below.

A large part of this book has been devoted to a phonological issue that is intimately related to our research question. Fairly recently some doubts have been cast on the universality of onset weightlessness, most notably by Davis (1985). He reported the existence of languages in which onsets did seem to influence stress locations. In some of these languages onsets seemed to contribute to weight directly, affecting stress locations by their mere presence, while in others only the prominence of some onsets exerted their influence on stress placement. In view of a universal rule that completely disregards onsets when calculating syllable weight, such languages are abominations. If the absence of weight in onsets is indeed a true universal, something else must be causing the observed stress patterns, at the same time creating the illusion of a quantity-sensitive stress rule that refers to onsets. Though the universality of onset weightlessness is not a holy wall that needs to remain unbreached, we have chosen to abide by the hypothesis that it remains unchallenged by these so-called “onset-sensitive languages”.[1] This position allows us to keep the number of possible quantity-sensitive stress rules fairly limited (refraining from an elaboration of the theoretical model for the sake of a few marginal cases only), but it forces us to closely inspect the offending languages to see whether we can find rules that generate the observed stress patterns while remaining within the accepted models of contemporary metrical phonology. We have taken up that task in the second part of this book, devoting one chapter to an overview of the onset-sensitive languages we know, and one chapter to a full analysis of two notorious cases. Summaries of these chapters can be found in section 7.2.2 below.

In the following section we will present an overview of the main results we obtained from the phonetic experiments and phonological investigations. Finally, in section 7.3, the general conclusions are drawn and some suggestions for further research are given.

7.2 Summary of main findings

7.2.1 Phonetics

In chapter 2 we assumed that the phonetic correlate of weight is duration. We departed from the hypothesis that this phonetic correlate should not be active in weightless units. The most straightforward clue for the inactivity of duration in onsets would be invariance of onset duration (i.e. whatever the number of segments, the duration remains the same). If that were the case, the weightlessness of onsets could easily be explained. No phonological contrast can be based on something that never changes.

---

[1] Note that, if onset weightlessness were not universal, the status of the auditory phenomenon underlying the weightlessness of onsets would be seriously affected. If some onsets may not be weightless, this phenomenon cannot be something that is of a general nature. Though we have found phonetic evidence for marginal onset influence on stress in our search for this phenomenon (see below), this influence does not lie in the domain of syllable weight, but rather in that of syllable prominence. Hence, we have no reason to believe that onset weight proper (i.e. the usage of true moraic onsets) is an option, and we may remain confident that this is caused by a general auditory phenomenon.

Since nothing in phonetics is ever that absolute, however, a more realistic prediction drawn from our hypothesis would be that onsets do vary in duration, though not as drastically as nuclei and codas. In other words, we expected onset duration to be *relatively invariant*. An asymmetry might exist between the durational behaviour of onset and coda clusters. Evidence for such an asymmetry can be found in Nooteboom & Cohen (1988) who present spectrograms of words with /s/, /st/ and /str/ onsets. The striking thing about these spectrograms is that the duration of the onset cluster remains largely the same irrespective of the number of consonants they contain. The coda clusters (/f/, /fs/ and /fst/) they present do increase in duration in accordance with segment number.

We decided to conduct a pilot experiment in which we checked whether the data presented in Nooteboom & Cohen were just isolated cases or whether they represented a structural property of onsets. In this production experiment we asked two speakers of Dutch to produce sentences that contained target words with onset and coda clusters with a variable number of segments. We used exactly those onset and coda clusters that Nooteboom & Cohen used for their spectrograms. We measured the average durations of these clusters and found that onset clusters are *not* durationally invariant. Their duration increases from roughly 135 ms for single segments to 175 ms for two segments and 225 ms for three segments. Coda durations increase less drastically from 100 ms to 160 ms and 190 ms, respectively. These duration increases by number of segments are all highly significant. We repeated the experiment with different target words in which onset and coda clusters were exact mirror images (/s/, /st/, /str/ - /rts/, /st/, /s/). The results were largely the same.

It appears that onset durations are not invariant, not even relatively: the increase in duration by number of segments is even larger in onsets than in codas. Even when we slightly changed the scope of the syllabic part that was to remain durationally invariant, and included everything from the beginning of the syllable to the intensity peak of the vowel (labelling this the *ascent*), we did not find durational invariance. We had to conclude, therefore, that an explanation for the weightlessness of the onset cannot be found in speech production.

In chapter 3 we conjectured that the absence of a durational asymmetry between onsets and codas in speech production was not a total loss, because we did not look at the perception side of the matter yet. If something that is present in the speech signal, say variable onset duration, cannot be perceived by the listener, then no phonological opposition, such as the difference between heavy and light syllables, may be based on it. Hence, if onset duration would be *perceptually invariant*, we could accept that as a valid reason for its weightlessness by the same token as when it had been invariant in speech production. Again we expected the invariance of onset duration to be relative, i.e. incorrect or inaccurate duration perception would be more characteristic of onsets than of nuclei and codas.

We tried to find evidence for the relative perceptual invariance of onset duration by changing the durations of monosegmental onsets, nuclei and codas in the Dutch words *mam* ‘mother’ and *sas* ‘good humour’. If onset duration is indeed perceptually invariant (relative to nucleus and coda duration) then the duration changes in onset segments should be perceived as less sizeable than such changes in nucleus and coda. So, we expected the *fidelity* with which duration differences are perceived to be better for nuclei and codas than for onsets (the actual duration differences in nuclei and codas are perceived more faithfully than duration differences in the onset).

In several adjustment experiments we asked listeners to reproduce the durations of reference stimuli with lengthened or shortened onsets, nuclei or codas (in either *sas* or *mam*) in an adjustable comparison stimulus, such that the two became perceptually equal. In the clearest results we obtained, a duration change of 30 ms in the onset of *sas* was reflected by an adjustment of the comparison stimulus of a mere 8 ms on average. The same 30-ms change in the coda segment resulted in the correct average adjustment of 29 ms, while the 30-ms change in the nucleus led to an exaggerated average of 59 ms for the adjusted signal. Clearly, the duration change in onsets is less well perceived than identical changes in nucleus and coda. We attributed the fact that duration differences were perceived as being longer for nuclei than for codas to the sonority difference between these segments. Further evidence for an influence of sonority on the perception of duration differences was found when we compared the results for the more sonorous *mam* to the results we found for *sas*. The overall perception of duration differences was more faithful for the former than for the latter.

Unfortunately, the results of the above experiments were inconclusive on the matter of *accuracy* in duration perception. We expected this accuracy to be poor for onsets, since duration changes in onsets are underestimated, which means that the duration of errors in the adjustment is underestimated (comparison signals with relatively large durational deviations from a reference signal are still judged to be durationally equal to this reference stimulus). As a result, we expect accuracy to be worse for onsets than it is for nuclei and codas. The traditional measure for the accuracy with which a particular property is perceived is the JND (Just Noticeable Difference). This JND is small if perception is accurate (even the slightest change is noticed). In our adjustment experiments we found no clear differences in the JNDs for onset, nucleus and coda duration (when caclculating JNDs from the standard deviations of the stimulus adjustments made by the subjects). Therefore we conducted a further experiment in which we asked subjects to perform a pairwise-comparison task which was better suited to yield clear JNDs. We composed stimulus pairs that consisted of one fixed reference stimulus of 300 ms combined with a comparison stimulus in which either onset, nucleus or coda duration was lengthened or shortened by 20, 40 or 60 ms. These pairs were offered to subjects who were asked to select the longer stimulus of the pair. The success with which subjects could perform this task reflected the JND for the segment that was altered in the comparison stimulus. Expressed as percentages of the reference stimulus duration the JNDs for nucleus, coda and onset we found are 6.6%, 10.6% and 16.2% respectively.

Thus, the poor perception of onset duration is reflected in the subjects' estimation of the magnitude of a duration change in these onsets as well as in the accuracy with which they perceive duration changes in onsets. We concluded, therefore, that the weightlessness of the onset is caused by the relatively poor performance of the listener when it comes to perceiving the phonetic correlate of weight in onsets. It seems indeed improbable that onset weight plays a role in language structure.

The answer to our research question that we have found in chapter 3 was not completely satisfactory, though. The question that immediately came to mind when we contemplated this answer was: *Why is duration perception in onsets so poor?* Before we answered that question we would not really know why onsets are weightless. In chapter 4, some psychophysical experiments were described that we conducted in our search for the reason that underlies the poor perception of onset duration.

We could think of several reasons why the duration of onsets should not be perceived as accurately as the durations of both nucleus and coda. As our first lead towards an answer we chose to adopt the view that there might be a certain trigger somewhere in the syllable that activates the mechanism with which we estimate its duration. In that case, the duration of all the material that comes before the trigger cannot be correctly estimated. So, if we could find evidence for the existence of such a trigger that occurs after the onset but before (most of) the vowel, we would have an explanation for the poor duration perception in onsets.

The trigger we were looking for was likely to be a salient point in the syllable (it seems logical that events can more easily be triggered by phenomena that can easily be perceived). Three candidates served as potential triggers: the intensity maximum, the P-centre (or perceptual moment of occurrence of the syllable), and the CV-transition or PIVOT. The design of the experiments needed to check which of these was the trigger we were looking for was simple. The location of a potential trigger needed to be varied in a syllable with a steady overall duration. In this way the portion of the syllable that came after the trigger would be varied. Since only the duration of that portion might be estimated correctly, the perceived duration of the syllable would have to decrease as the trigger moved to the end of the syllable.

In our first psychophysical experiment we assumed that the intensity maximum was the trigger. We generated syllabic stimuli (*mam*) in which the location of the intensity maximum was varied over almost the entire length of the syllable. We tested whether variable intensity peak location led to significant differences in the perceived durations of these stimuli in a pairwise- comparison experiment and in a duration estimation experiment. Both these experiments yielded no effects. The subjects judged all the stimuli to be equally long.

With respect to the second trigger, the P-centre, the literature shows that the changes in the location of the intensity peak should not be without effect. According to Pompino-Marschall (1989) the location of the P-centre depends on the location of the intensity peak. Hence, the stimuli we used in the intensity experiment should have variable P-centres.[2] Since we already knew that these stimuli did not have perceptually different durations, we only needed to prove that P-centres in these stimuli were indeed variable to discard the P-centre as a trigger candidate. In a rhythm based experiment the P-centre locations of those stimuli were determined. It was found that they moved to the right edge of the syllable in harmony with the location of the intensity peak. Since these differences in P-centre location did not lead to differences in perceived duration, the P-centre could not be the trigger.

We were not able to decide whether the third candidate, the CV-transition, was the trigger we were looking for. It was by definition impossible to vary the position of the CV-transition in an experiment without altering the durations of onset and nucleus. In the perception experiments described in chapter 3 we had altered these onset and nucleus durations and, consequently, changed the relative position of the CV-transition. We knew that this led to differences in the perceived duration, but we could not be sure whether this was due to the variable trigger location or some other effect of the altered onset duration. For

[2] This was an unavoidable result. However, if the intensity experiment would have yielded positive results we could have determined later whether these were due to P-centre or to intensity peak variations.

lack of any conclusive evidence that the PIVOT is indeed the trigger, we felt forced to direct our attention elsewhere, keeping the PIVOT-trigger hypothesis in the back of our minds.

In our final experiment we tried to determine whether there was a difference in the perception of duration differences in the constituents of speech and non-speech stimuli. The duration asymmetry we have found could be due to a phonological distinction that is specific to speech. The automatic consequence would then be that duration perception in non-speech signals made up of several different parts would be equal for all those parts, since the phonological distinction introduced above is not relevant when we listen to non-speech. For this final experiment we composed a non-speech "syllable" that consisted of a periodic sawtooth signal (nucleus) flanked by two noise parts (onset and coda). With this artificial syllable we repeated the pairwise-comparison experiment from chapter 3. No differences in duration perception were found for the three constituents of the non-speech syllable (for discussion see section 7.3).

With the conclusions from all our experiments in mind, we were able to propose a tentative model of the syllable that is operative in weighting processes. We adopted a prosodic syllable of which the normal (machine observable) intensity curve was "pushed up" by moraic building blocks, but only in the rhyme part of the syllable (nucleus in all cases and coda optionally). The combination of sonority and moraicity thus created was labelled *prominence* and claimed to be the relevant syllabic property in weighting processes. The general rule was: heavy syllables have more area under the prominence curve than light syllables. To make life easier we assumed that heavy/light oppositions are binary and, for syllables to be heavy, their total prominence has to exceed a certain threshold. This threshold was represented by a horizontal line through the prosodic syllable. Apparently, languages may differ in the height at which they place the threshold (or: in the amount of prominence needed to make a syllable heavy). In most languages the threshold is placed such that the difference between heavy and light syllables is made solely on the basis of the huge contribution that moraic segments make to syllabic prominence. However, some languages exist that place the threshold so high that more than moraicity is needed to make syllables heavy. In some cases the sonority added by sonorant codas is sufficient, in others only syllables with the most sonorant vowels are prominent enough to exceed the threshold. We have expressed the belief that only prosodically active properties, like duration, pitch and sonority, may feature in stress rules, since these are the only ones that can alter the prominence curve. Recognition of non-moraic contributions to the total prominence automatically means that such contributions of onsets may also play a role. This is an unlikely event, though, since the prominence threshold would have to be extremely low for onsets to matter. Yet, some of these onset prominence stress systems seem to exist (see section 7.2.2).

7.2.2 Phonology

As was noted above, the existence of languages was reported in which onsets were claimed to contribute to weight. Our own research has revealed one Amazonian language and 14 Australian languages in which the onset seems to be moraic: it contributes to weight whenever it is present, making onsetless syllables light and all others heavy. Furthermore, 3 Australian, one Micronesian and one Amazonian language could be found in which the location of main stress seemed to be influenced by the identity, not the mere presence, of the syllable onset. Together, these languages form only a small proportion of the stress languages in the world. However, the group is large enough, and the problem these languages present to metrical theory is serious enough, to warrant closer investigation.

Though we were inconclusive as to the cause of it, our experiments did reveal a durational asymmetry between onsets and codas. Our main assumption was that this asymmetry reflects the phonological weightlessness of the onset. Hence, we had to conclude that onsets cannot be moraic. Thus, like the phonologists who claim onset weightlessness merely on the basis of observation, we have a problem with the so called onset-sensitive stress languages that seem to have moraic onsets (the first group: 15 in total). In chapter 5, after an introduction to the rules that are used in metrical theory (which was at the same time an overview of the types of stress system that can be found on the Australian continent), we tackled this problem.

We selected two Australian languages as representatives for the group of 15. These languages were Western Aranda and Alyawarra, both spoken in Central Australia.[3] It was observed that the languages in question only use the difference between heavy and light syllables at the beginning of words. In fact initial syllables were the only ones that could be light. In words that had such an initial light syllable, main stress was placed on the second syllable, while it was initial in all other cases. We proposed that this odd distribution of light syllables, and the resulting odd stress pattern, are the result of a historical phenomenon called Initial Dropping.

[3] These were the two best described languages in the group. Most of the other languages in this group are (or were) spoken on the Cape York peninsula in Queensland, to the northeast of Aranda and Alyawarra.

This process, which occurred without exception in all the Aboriginal languages we considered, involved the loss of the first consonant of all words in the language, often taking with it the following vowel if that was short. In case the following vowel was long, Initial Dropping reduced it to a short vowel. We further claimed that this resulting onsetless initial syllable was not a proper stress bearer in these languages (in which nearly *all* other syllables had onsets). This resulted in the movement of stress to the second syllable in words that had such an initial syllable. Of course, in words of which the short vowel in the first syllable was deleted the stress obligatorily moved to the second syllable. Yet, in these cases the initial syllable had become completely invisible, so that the stressed second syllable seemed to be initial. Thus, the observed stress pattern was created.

We proposed to derive this pattern in metrical theory by letting the metrical rules skip the initial vowels somehow. Several skipping methods were reviewed. Our own alternative was presented in chapter 6 for Western Aranda and Alyawarra. We proposed to skip the initial vowels through misalignment of the initial foot in an Optimality Theoretic framework. This misalignment can be achieved if one assumes that the trochaic feet, which we use to analyse the stress patterns of all these languages, must begin with an onset. The first onset of vowel-initial words is that of the second syllable, so the (left-headed) trochee starts there with a stress, skipping the onsetless initial syllable. In the other cases the trochee starts at the word edge (at which we find an onset) deriving initial stress. This misalignment approach was claimed to cover all the cases for which we seemed to need moraic onsets. Small differences, like the location of stress in disyllabic words, could easily be covered by differences in ranking of the Optimality Theoretic constraints we used to analyse these stress patterns, as was shown in chapter 6.

Two languages from the second group of onset-sensitive stress systems that was introduced in chapter 5 could also be analysed without moraic onsets. In these cases (Puluwat and Pirahã) the possibility of onset prominence that we hinted upon above was exploited. The prominence threshold in these languages is placed such that the sonority of the vowel *with respect to* the onset is decisive in the categorisation of the syllable as heavy or light. We have noted that the two languages in this second group do not form very strong cases for onset-sensitivity since their stress rules are not undisputed.

The remaining cases from the second group formed a third group that remained problematic. In the languages from this group (which were all Australian) a non-prosodically active segmental property seemed to influence the location of stress. We considered influence of non-prosodic features on prominence, and thus stress, to be impossible and predicted the existence of other rules that conspire to generate the correct stress pattern and explain the correlation between the segmental property and the location of stress. One such case is Mathimathi. Main stress in this language falls on the second of two light syllables at the left word edge if that second syllable begins with a coronal onset or vowel in hiatus. Since the feature [coronal] is not prosodic (and vowels in hiatus have no onset at all) we could not simply let coronal onsets contribute to prominence while ignoring non-coronals. In the final part of chapter 6 we showed, with the help of Gahl (1996), that the Mathimathi stress pattern was the result of morphologically conditioned borrowing of a stress shift rule that seemed to have been an areal feature of the region in which Mathimathi was spoken. In all words that had disyllabic stems, stress shifted to the second syllable, while stress remained initial in words that had monosyllabic stems. Furthermore, medial consonants in disyllabic stems lenited to coronals or were deleted. Thus, the observed stress pattern arose. In the final section of chapter 6, this solution to the Mathimathi problem was cast into an Optimality Theoretic analysis.

7.3 Final conclusions and suggestions for further research

In conclusion we can claim with reasonable certainty that the cause of onset weightlessness cannot be found in the actual physical durations of the segments contained in onsets compared to the durations of nucleus and coda segments. Though production experiments on segment duration can reveal differences between heavy and light syllables (Duanmu 1994) or even the difference between moraic and non-moraic codas (Broselow, Chen & Huffman 1997), we claim that the intrinsic weightlessness of onsets is a phenomenon of a different order, since it is universal; no comparison between moraic and non-moraic onsets can ever be made. For this obligatory weightlessness, no evidence can be found in speech production. If anything, the durational behaviour of onsets in speech production could more easily be interpreted as evidence for onset weight and nucleus and coda weightlessness. As has been observed in the past, and again in chapter 2, the effect of duration increases in nuclei or codas is often compensated by a (relative) duration decrease in the other part of the rhyme (for example, if the coda lengthens, the nucleus shortens). Hence, the total “gain” in syllable duration caused by a duration increase in nuclei or codas is relatively small. Such compensatory effects are largely absent when we lengthen the onset. Any duration increase of the onset translates to an almost as sizeable increase in total syllable duration.

Therefore, if actual physical duration would be anything to go by, the stable and relatively independent durational behaviour of onsets would make them prime candidates for moraicity rather than weightlessness. However, the existence of moraic onsets is not in agreement with the linguistic facts. Languages that use onset weight cannot be found. In chapters 5 and 6 we have seen that the languages for which we did seem to need onset weight appeared to be totally non-moraic under closer scrutiny. No language we know of refers to moraic onsets in the distinction between heavy and light syllables. Coda and nucleus weight, on the contrary, are amply used. Therefore, we must assume that physical duration is not the phonetic correlate of phonological weight and that the possible relation between the absence of compensatory effects for onsets and weight given above is meaningless.

The evidence from the perception experiments in chapter 3 shows that duration *perception* and phonological weight are related. Duration perception is relatively poor in weightless segments when compared to duration perception in segments that have weight. The fact that this difference is not made in a related non-speech signal, as is shown in chapter 4, could force us to assume that onset weightlessness is an abstract phenomenon. The observed asymmetry in the perception of their durations might be something that is forced upon us by the linguistic centre in our brain when we listen to speech. For some, still unknown, reason the abstract syllabic model in our heads may contain no weight unit for onsets. Since the presence of such a unit enhances duration perception, the difference between onsets, nuclei and codas falls out automatically, reflecting the resulting phonetic correlate of this abstract phenomenon. In the discussion at the end of section 4.4.3 we have noted, however, that we must be cautious in drawing this conclusion. A likely alternative is still that the PIVOT is the trigger of faithful duration perception. It may well have been the case that some crucial PIVOT property was not present in the artificial syllable we used in the non-speech experiment. If that is the case, it is only logical that we did not find differences in the duration perception of artificial “onsets”, “nuclei” and “codas”. If the absence of an effect for non-speech is indeed caused by our experimental design, new experiments must be conducted to find out whether the asymmetry between speech and non-speech does exist. Before we can draw the linguistic conclusion presented above we must make sure that we have used a representative artificial syllable for the non-speech experiment.

At the end of section 4.4.3 we have suggested several ways in which we could improve the artificial syllable such that it might have PIVOT-like properties at the “CV-transition”. Changing the identity of the "onset" signal from noise to something that sounds more like the sawtooth in the "nucleus" might introduce the simulation of a rapid change of the first formant (related to mouth opening), which might be the triggering property of the PIVOT. A completely different artificial syllable that we might compose is one in which we use the method of sine-wave analogons (see Parker 1988). In this method the formants in the speech stream are replaced by sine waves of the formant's frequency, creating F0-less "speech" with discrete formant values tracing the formant tracks that could be found in the speech signal. It might be that exactly those properties of speech that are necessary to trigger duration perception differences are retained in the sine-wave analogon stimulus. A rapid change in the formants, for instance, can still be detected. Hence, a PIVOT equivalent is still present in the signal, which may act as a trigger for faithful duration perception. Possibly these non-speech stimuli will reveal the durational asymmetries we have missed with our noise-sawtooth-noise stimulus. If so, the cause of onset weightlessness must be psychophysical, and not linguistic, in nature. That would mean we must take a step back to where we were at the beginning of section 4.4. We would have to take it from there and search for a way to check whether the PIVOT does indeed act as a trigger for faithful duration perception, or whether another, as yet undiscovered, property of onsets in speech and non-speech causes weightlessness. In the latter case, conducting further experiments in psychophysics or neurology may be the way to go about finding an answer to our research question (cf. section 4.3.2). Respecting the relative importance of the excursion size in the CV transition that we have discovered in the stress rules for Puluwat and Pirahã, we might run follow-up experiments in which we systematically alter the properties of the PIVOT itself. If, for instance, a more sharply defined PIVOT leads to more accurate duration preception of the syllabic part after it, we have found evidence for the PIVOT as a duration perception trigger.

If the follow-up experiments yield a persistent difference in the duration perception of speech and non-speech we must assume that onset weightlessness is linguistically determined. That would mean we have found only half an answer to our research question. We would know then that the syllable onset is weightless because the syllabic model that is part of our internal grammar forces it to be, and we would have have found the phonetic correlate (perceived duration differences) that indicates the difference between onsets and codas. However, we would not have an answer to the real question that follows from our discovery: *Why does the abstract syllable model in our heads feature weightless onsets?* We do not really know an answer, but we suggest some ideas that might lead to further research into what will then be a psycholinguistic phenomenon.

First of all we could check whether the claim holds for other languages. Had onset weightlessness been caused by a general psychophysical phenomenon, we might have expected speakers of other languages to perform similar to the Dutch subjects in the experiments described above. Speakers of quantity-insensitive languages might do somewhat worse, but the generality of the phenomenon that would have caused the weight related difference between onsets and codas should have made sure that we were to find an effect in all cases. We would have simply concluded then that some languages employ this phenomenon present in the speech signal to make phonological weight differences, and others do not.

In case the difference is linguistic, however, speakers of languages that do not have moraic codas, for instance, may perceive duration differences in codas as poorly as they do perceive duration differences in onsets. A repetition of the pairwise-comparison experiment described at the end of chapter 3 with speakers of a language that only has moraic vowels should only yield a duration perception asymmetry between these vowels on the one hand and onsets and codas on the other. This asymmetry should be larger than the difference between the nucleus and the other two constituents that we are to find should we repeat the experiment with speakers of a language that is quantity-insensitive altogether. In that case only the sonority difference between vowels and consonants can contribute to the differences in duration perception. As a final test we could check with speakers of a language in which the sonority of (onset) consonants is decisive for their weight, whether these subjects are more sensitive to duration differences in sonorant segments (with respect to non-sonorant segments) than our Dutch listeners were. In this respect, an experiment with speakers of a language for which onsets were claimed to be moraic could also yield some interesting results.

If all these experiments yield positive evidence for the claim that onset weightlessness is phonologically determined we can start looking for the reason behind it. One possibility we can think of is theory internal. Given the binarity hypothesis that underlies most aspects of generative grammar, we might assume that only two moras can maximally be assigned to one syllable. If a second mora is assigned to a syllable it goes either to the nucleus itself or its dependent, the coda. We do not know the reason for this choice, but if we can find it we might have a lead to why the onset is weightless. The result of this choice is namely that, if we were to assign a mora to an onset, it would most likely be the third, which results in a trimoraic syllable. Obligatory onset weightlessness might be the grammar’s first and foremost safeguard against these trimoraic syllables. Many versions of moraic theory disallow those trimoraic syllables, yet others do allow trimoraic rhymes. Further theoretical (and possibly experimental) research is needed to verify that syllables can be maximally bimoraic. If so, binarity may provide a theory internal reason for the obligatory weightlessness of the onset.

Still other ways may be thought of that help us to find a phenomenon that takes away our attention from onsets when we consider the weight of syllables. The discovery of this phenomenon may well contribute to the better understanding of what linguistics in general and stress in particular are all about.

Some final remarks concern the model for the prosodic syllable that we have developed on the basis of our experimental results. This model depends heavily on language research. The observations we made in chapter 3 with respect to the influence of moras and sonority on duration perception (and hence, possibly on stress rules) could easily be corroborated by language data. Using these observations we generated a model that could be further tested against language data. However, some of the languages that use stress rules relevant to the structure of our model were not studied very extensively, and there may be other relevant languages that we have not yet discovered. In-depth studies of these languages may reveal facts on the basis of which our model needs to be changed. For instance, if Pirahã and Puluwat appear to be onset-insensitive at closer scrutiny, or if there are Mathimahti like rules for these languages, with which we can derive the stress patterns without reference to onset prominence, we can eliminate *all* influence of onsets from our model. That could mean that the usage of prominence depends on moraicity, and that, if a segment is not moraic, it cannot play a role in prominence-sensitive stress rules.

For Djapu we still need to find a stress rule that circumvents the usage of non-prosodically active features in a prominence-sensitive stress rule. If we cannot find such a rule, and for instance, place features are found to irrefutably influence stress assignment, we must give up the claim that the prominence area in our model only represents prosodically active features. However, as we have noted earlier, the fact that we were able to find an alternative for Mathimathi and Ngarigu, in which non-prosodic prominence does not play a role, strengthens our belief that prominence must be prosodic and that an alternative can be found for Djapu as well.

Of course, we expect that any language for which onset-sensitivity may be claimed in the future will not use truly moraic onsets. It will be either an onset prominence system that abides by our prosodic laws, or a Mathimathi-like case in which prominence-sensitivity is only apparent.

References

Abel, S.M. (1972). Duration discrimmination of noise and tone bursts. *JASA* 51, 1219–1223.

Alpher, B. (1976). Phonological peculiarities of Yir Yoront song words. In: P. Sutton (ed.) *The languages of Cape York*, 78–83. AIATSIS, Canberra.

Anderson, J. (1969). Syllabic or non-syllabic phonology. *JL* 5, 136–143.

Archangeli, D.B. (1986). Extrametricality and the percolation convention. Unpublished Ms, Dept. of Linguistics, University of Arizona.

Beckman, M.E. (1986). *Stress and non-stress accent*. Netherlands Phonetic Archives 7. Foris, Dordrecht.

Beckman, M.E. & J. Edwards (1990). Lengthenings and shortenings and the nature of prosodic constituency. In: J. Kinston & M.E. Beckman (eds.). *Papers in laboratory phonology I: Between the grammar and the physics of speech*, 152–178. Cambridge University Press, Cambridge.

Berinstein, A. E. (1979). A cross-linguistic study on the perception and production of stress. *UCLA working papers in phonetics* 47. Dept. of Linguistics, UCLA.

Berry, L. (1996). Alignment and adjacency in Optimality Theory: Evidence from Warlpiri and Arrernte. PhD Dissertation, University of Sydney.

Bezooijen, R. van & L.C.W. Pols (1993). Evaluation of text-to-speech conversion for Dutch. In: V.J. van Heuven & L.C.W. Pols (eds.) *Analysis and synthesis of speech: Strategic research towards high-quality text-to-speech generation*, 339–360. Mouton de Gruyter, Berlin.

Blevins, J. (1995). The syllable in phonological theory. In: J. Goldsmith (ed.). *The handbook of phonological theory*, 206–244. Blackwell Publishers, Cambridge, Massachusetts.

Bolinger, D. (1958). A theory of pitch accents in English. *Word* 14, 109–149.

Bolinger, D. (1972). Accent is predictable (if you're a mind-reader). *Language* 48, 633–644.

Breen, G. & R. Pensalfini (forthcoming). Arrernte: A language with no syllable onsets. To appear in Linguistic Inquiry.

Bromberger, S. & M. Halle (1988). Conceptual issues in morphology. Ms, MIT.

Broselow, E. (1995). Skeletal positions and moras. In: J. Goldsmith (ed.). *The handbook of phonological theory*, 175–205. Blackwell Publishers, Cambridge, Massachusetts.
Broselow, E., S. Chen & M. Huffman (1997). Syllable weight: Convergence of phonology and phonetics. *Phonology* 14:1, 47–82.
Browman, C. & L. Goldstein (1988). Some notes on syllable structure in articulatory phonology. *Phonetica* 45, 140–155.
Buller, B., E. Buller & D. Everett (1993). Stress placement, syllable structure and minimality in Banawá. *IJAL* 59, 280–293.
Bunn, G. & R. Bunn (1970). Golin phonology. *Pacific Linguistics*, series A-23, 1–7. Canberra.
Burghardt, H. (1973). Die subjective Dauer schmallbandiger Schall bei verschiedenen Frequenzlagen. [The subjective duration of narrow band noise at different frequencies] *Acustica* 28, 278–284.
Cambier-Langeveld, T. (1997). The domain of final lengthening in the production of Dutch. In: J. Coerts & H. de Hoop (eds.) *Linguistics in the Netherlands 1997*, 13–24. John Benjamins, Amsterdam.
Cardozo, B.L. (1965). Adjusting the method of adjustment: SD vs SL. *JASA* 37, 786–792.
Caspers, J. (1994). Pitch movements under time pressure. HIL PhD Dissertations 10. Holland Academic Graphics, The Hague.
Chen, M. (1970). Vowel length variation as a function of the consonant environment. *Phonetica* 22, 129–150.
Chomsky, N. & M. Halle (1968). *The sound pattern of English*. Harper and Row, New York.
Churchyard, H. (1991). Compensatory lengthening and “gemination throwback” in Trukese and Puluwat as evidence for the rime in moraic theory. Paper presented at LSA meeting, Chicago. Ms, University of Texas.
Clements, G.N. (1990). The role of the sonority cycle in core syllabification. *Papers in laboratory phonology I: Between the grammar and the physics of speech*, 283–333. Cambridge University Press, Cambridge.
Clements, G.N. & S. Hertz (1996). An integrated approach to phonology and phonetics. In: B. Laks & J. Durand (eds.) *Current trends in phonology: Models and methods*, 143–174. European Studies Research Institute, University of Salford.
Crystal, T.H. & A.S. House (1988). Segmental durations in connected speech signals: Current results. *JASA* 83, 1553–1573.
Dainora, A. & E. Steinberg (1996). Rethinking onset weight in the sonority model of the syllable. Paper presented at the Eighth International Phonology meeting in Vienna.

Davis, S. (1985). *Topics in syllable geometry*. PhD Dissertation, University of Arizona.
Davis, S. (1988). Syllable onsets as a factor in stress rules. *Phonology* 5:1, 1-19.
Davis, S. (1995). Syllable stucture & moraic phonology. Lecture notes from the 1995 Linguistic Institute in New Mexico.
Dixon, R.M.W. (1970) The languages of the Cairns Rainforest region. *Australian linguistic studies in honour of Arthur Capell*. Pacific Linguistics Series C-13. Canberra.
Dixon, R.M.W. (1980). *The languages of Australia*. Cambridge University Press, Cambridge.
Downing, L. (1993). Unsyllabified vowels in Aranda. *CLS* 29, 171–185.
Downing, L. (1998). On the prosodic misalignment of onsetless syllables. *NLLT* 16:1, 1–52.
Dogil, G. (to appear). The phonetic manifestation of word stress. To appear in: H. van der Hulst (ed.) Word prosodic systems in the languages of Europe. Mouton de Gruyter, Berlin.
Dogil, G. & G. Braun (1988). *The PIVOT model of speech parsing*. Academie der Wissenschaften, Wien.
Dogil, G. & H.C. Luschuetzky (1990). Notes on sonority and segmental strength. *Revista di Linguistica* 2:2, 3–54.
Duanmu, S. (1994). Syllabic weight and syllabic duration: A correlation between phonology and phonetics. *Phonology* 11, 1–24.
Elbert, S.H. (1972) Puluwat dictionary. *Pacific Linguistics*, series C-24. Canberra.
Everett, D. (1988). On metrical constituent structure in Pirahã. *Natural Language and Linguistic Theory* 6:1, 207–246.
Everett, D. (1995). Quantity, sonority and alignment constraints in Suruwahá and Banawá prosody. Ms, University of Pittsburgh.
Everett, D. & K. Everett (1984). Syllable onsets and stress placement in Pirahã. *Proceedings of the West Coast Conference on Formal Linguistics* 3, 105–116.
Fowler, C.A. (1979). ‘Perceptual centers’ in speech production and perception. *Perception & Psychophysics* 25, 375–388.
Fry, D.B. (1955). Duration and intensity as physical correlates of linguistic stress. *JASA* 27, 765–768.
Fry, D.B. (1958). Experiments in the perception of stress. *Language and Speech* 1, 126–152.
Fry, D.B. (1965). The dependence of stress judgements on vowel formant structure. In: *Proc. of the 5th ICPhS, Munster 1964*, 306–311. S. Karger, Basel/ New York.
Fudge, E. (1969). Syllables. *JL* 5, 253–287.

Fujisaki, H., K. Nakamura & T. Imoto (1975). Auditory perception of duration of speech and non-speech. In: G. Fant & M.A.A. Tatham (eds.) *Auditory Analysis and Perception of Speech*, 197–219. Academic Press, London/New York/San Francisco.

Gahl, S. (1996). Syllables as a factor in stress rules: The case of Mathimathi revisited. *Phonology* 13:3, 329–344.

Gimson, A.C. (1975). *An introduction to the pronunciation of English*. Edward Arnold, London.

Goedemans, R.W.N. (1993). Rhymes & riddles: Investigations in the phonetics and phonology of the weightlessness of syllable onsets. MA Thesis, Leiden University.

Goedemans, R.W.N. (1996a). Syllable weight and prominence. In: R.W.N. Goedemans, E.A.M. Visch & H.G. van der Hulst (eds.), *Stress patterns of the world: Part 1, Background*, 115–163. HIL Publications 2. Holland Academic Graphics, The Hague.

Goedemans, R.W.N. (1996b). An optimality account of onset-sensitive stress in quantity-insensitive languages. *The Linguistic Review* 13:1, 33–48.

Goedemans, R.W.N. (1997). Putting the Mathimathi stress rule in its proper perspective. *Australian Journal of Linguistics* 17, 43–69.

Goedemans, R.W.N. & V.J. van Heuven (1993). A perceptual explanation of the weightlessness of the syllable onset. *Proceedings Eurospeech '93*, Berlin, 1515–1518.

Goedemans, R.W.N. & V.J. van Heuven (1995). Duration Perception in Subsyllabic Constituents. *Proceedings Eurospeech '95* Madrid, 1315–1318.

Goedemans, R.W.N., H.G. van der Hulst & E.A.M. Visch (1996). *Stress patterns of the world: Part 1, Background.* HIL Publications 2. Holland Academic Graphics, the Hague.

Goedemans, R.W.N. (in prep). Aboriginal stress: An overview. Ms, ANU and Leiden University.

Goldsmith, J. (1976). Autosegmental phonology. PhD Dissertation, MIT, Cambridge, Massachusetts (distributed by the Indiana University Linguistics Club, Bloomington, Indiana).

Gussenhoven, C. & A.C.M. Rietveld (1992). Intonation contours, prosodic structure and preboundary lengthening. *Journal of Phonetics* 20, 283–303.

Halle, M. & J.-R. Vergnaud (1980). Three dimensional phonology. *Journal of Linguistic Research* 1, 83-105.

Halle, M. & J.-R. Vergnaud (1987). *An essay on stress*. MIT Press, Cambridge, Ma.

Handel, S. (1989). *Listening*. The MIT Press, Cambridge Ma.

Hayes, B.P. (1979). Extrametricality. *MIT Working Papers in Linguistics* 1, 77–86.

Hayes, B.P. (1985). Iambic and trochaic rhythm in stress rules. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 11, 429-446.

Hayes, B.P. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20, 253–306.

Hayes, B.P. (1995). *Metrical stress theory: Principles and case studies*. The University of Chicago Press, Chicago.

Hercus, L.A. (1982). *The Bāgandji Language*. Pacific Linguistics, Series B-67. Canberra.

Hercus, L.A. (1986). *Victorian languages: A late survey*. Pacific Linguistics, Series B-77. Canberra.

Hercus, L.A. (1989). Three linguistic studies from far south-western NSW. *Aboriginal History* 13:1, 44–62.

Hermes, D.J. & J.C. Gestel (1991). The frequency scale of speech intonation. *JASA* 90, 97–102.

Heuven, V.J. van (1987). Stress patterns in Dutch (compound) adjectives: Acoustic measurements and perception data. *Phonetica* 44, 1–12.

Heuven, V.J. van. (1994). What is the smallest prosodic domain? In: P.A. Keating (ed.) *Papers in Laboratory Phonology III, between the grammar and physics of speech*, 76–98. Cambridge University Press, Cambridge.

Heuven, V.J. van & M.P.R. van den Broecke (1982). Interactions of spectral distribution and decay time in the auditory perception of duration and their relevance to speech. *Progress Report of the Institute of Phonetics Utrecht* 7:1, 21–41.

Heuven, V.J. van & A.M.C. Sluijter (1996). Notes of the phonetics of word prosody. In: R.W.N. Goedemans, E.A.M. Visch & H.G. van der Hulst (eds.) *Stress patterns of the world: Part 1, Background*, 233–269. HIL Publications 2. Holland Academic Graphics, The Hague.

Hockett, C.F. (1955). *A manual of phonology*. Indiana University Press, Bloomington.

Hofhuis, E. (1993). Establishing prosodic structure by measuring segment duration. In: D. House & P. Touati (eds.) *Lund University Working Papers* 41, 136–139.

Hooper, J. (1976). *An introduction to natural generative phonology*. Academic Press, New York.

Howell, P. (1988). Prediction of P-center location from the distribution of energy in the amplitude envelope I,II. *Perception & Psychophysics* 43, 90–93, 99.

Hulst, H.G. van der (1984). *Syllable structure and stress in Dutch.* Foris, Dordrecht.

Hulst, H.G. van der (1996a). Separating primary and secondary accent. In: R.W.N. Goedemans, E.A.M. Visch & H.G. van der Hulst (eds.) *Stress patterns of the world: Part 1, Background*, 1–25. HIL Publications 2. Holland Academic Graphics, the Hague.

Hulst, H.G. van der (1996b). Primary accent is non-metrical. *Revista di Linguistica* 9:1, 91–119.

Hulst, H.G. van der & G. Rowicka (1997). On some parallels between (unrealized) empty nuclei and (unstressed) syllables. Ms, HIL/Leiden University.

Hyman, L. (1985). *A theory of phonological weight*. Publications in Language Sciences 19, Foris, Dordrecht.

Idsardi, W. (1992). The computation of prosody. PhD Dissertation, MIT.

Ito, J. (1989). A prosodic theory of epenthesis. *Natural Language and Linguistic Theory* 7, 217–259.

Jespersen, O. (1904). *Elementarbuch der Phonetik.* B. Teubner, Leipzig.

Jones, D. (1950). *The phoneme: Its nature and use*. Heffner, Cambridge.

Jones, D. (1957). *An outline of English phonetics*. Heffner, Cambridge.

Jongenburger, W. & V.J. van Heuven (1991). The distribution of word-initial glottal stop in Dutch. In: F. Drijkoningen, A. van Kemenade (eds.) *Linguistics in the Netherlands 1991*, 101–110. John Benjamins, Amsterdam.

Kahn, D. (1976). *Syllable-based generalizations in English phonology*. PhD Dissertation, MIT, Cambridge, Massachusetts (distributed by the Indiana University Linguistics Club, Bloomington, Indiana).

Kalstrom, M.R. & E.V. Pike (1968). Stress in the phonological system of Eastern Popoloca. *Phonetica* 18, 16–30.

Kenstowicz, M. (1994). Sonority driven stress. Ms, MIT.

Kiparsky, P. (1991). Catalexis. Ms, Stanford University and Wissenschaftskolleg zu Berlin.

Klatt, D.H. & W.E. Cooper (1975). Perception of segment duration in sentence contexts. In: A. Cohen & S.G. Nooteboom (eds.) *Structure and Process in Speech Perception*, 69–89. Springer-Verlag, Berlin/Heidelberg/New York.

Kučera, H. (1961). *The Phonology of Czech*. Mouton, The Hague.

Ladd, R. (1980). *The structure of intonational meaning: Evidence from English.* Indiana University Press, Bloomington.

Ladefoged, P. (1982). *A course in phonetics*. Harcourt Brace Jovanovich, San Diego, 3rd Edition 1993.

Lahiri, A. & J. Koreman (1988). Syllable weight and quantity in Dutch. *Proceedings of the west coast conference on formal linguistics* 7, 217–228.

Langeweg, S.J. (1988). *The stress system of Dutch*. PhD Dissertation, Leiden University.

Lass, R. (1984). *Phonology: An introduction to basic concepts*. Cambridge University Press, Cambridge.

Laziczius, J. (1961). *Lehrbuch der Phonetik*. Akademie-Verlag, Berlin.

Lefkowitz, N.J. (1987). Talking backwards and looking forwards: The French language game Verlan. PhD Dissertation, University of Washington.

Lehiste, I. (1960). An acoustic phonetic study of internal open juncture. *Supplement to Phonetica* 5, 1–54.

Lehiste, I. (1970). *Suprasegmentals*. The MIT Press, Cambridge, Ma.

Lehiste, I. (1971). Temporal organization of spoken language. In: L. Hammerich, R. Jacobson & E. Zeirner (eds.) *Form and substance: Phonetic and linguistic papers presented to Eli Fischer-Jorgensen*, 159–169. Akademisk Forlag, Denmark.

Lehiste, I. (1980). Phonetic manifestation of syntactic structure in English. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo* 14, 1–27.

Levinsohn, S.H. (1976). *The Inga language*. Janua Linguarum, Series Practica 188. Mouton, The Hague.

Liberman, M. (1975). The intonational system of English. PhD Dissertation, MIT, Cambridge, Massachusetts (distributed by the Indiana University Linguistics Club, Bloomington, Indiana).

Liberman, M. & A.S. Prince (1977). On stress and linguistic rhythm. *Linguistic Inquiry* 8, 249–336.

Lindblom, B., B. Lyberg & K. Holmgren (1981). Durational patterns of Swedish phonology: Do they reflect short-term motor memory processes? Ms, distributed by the Indiana University Linguistics Club. Bloomington, Indiana.

Marcus, S.M. (1981). Acoustic determinants of Perceptual center (P-center) location. *Perception & Psychophysics* 30, 122–140.

McCarthy, J.J. & A. Prince (1986). Prosodic morphology. Ms, University of Massachusetts, Amherst, and Brandeis University.

McCarthy, J.J. & A. Prince (1993). Generalized alignment. Ms, University of Massachussets, Amherst, and Rutgers University.

McDonald, M. (1977). A study of the phonetics and phonology of Yaraldi and associated dialects. PhD Dissertation, ANU.

Moorhouse, M. (1846). *A vocabulary and outline of the grammatical structure of the Murray River language, spoken by the natives of South Australia from Wellington on the Murray as far as the Rufus.* Andrew Murray, Adelaide.

Morton, J., S.M. Marcus & C.R. Frankish (1976). Perceptual centers (P-centers). *Psychological Review* 83, 405–408.

Nooteboom, S.G. (1972). Production and perception of vowel duration: A study of durational properties of vowels in dutch. PhD Dissertation, Utrecht University.

Nooteboom, S.G. & A. Cohen (1988). *Spreken en verstaan.* [Speaking and understanding] Van Gorkum, Assen/Maastricht.

Nooteboom, S.G. & G.J.N. Doodeman (1980). Production and perception of vowel length in spoken sentences. *JASA* 67:1, 276–287.

Ohala, J.J & H. Kawasaki (1986). Prosodic phonology and phonetics. *Phonology* 3, 113–127.

Parker. E.M. (1988). Auditory constraints on the perception of voice-onset time: The influence of lower tone frequency on judgments of tone-onset simultaneity. *JASA* 83:4 1597–1607.

Payne, J. (1990). Asheninca stress patterns. In: D.L. Payne (ed.) *Amazonian Linguistics*, 185–209. University of Texas Press, Austin.

Pensalfini, R. (1996). Arrernte syllabification. In: J. Costa, R. Goedemans & R. van de Vijver (eds.) *ConSole IV Proceedings*, 237–249. SOLE, Leiden.

Perlmutter, D. (1995). Phonological quantity and multiple association. In: J. Goldsmith (ed.) *The handbook of phonological theory*, 307–317. Blackwell Publishers, Cambridge Massachusetts.

Peterson, G.E. & I. Lehiste (1960). Duration of syllable nuclei in English. *JASA* 32, 693–703.

Pike, K. & E. Pike (1947). Immediate constituents of Mazatec syllables. *IJAL* 13, 78–91.

Pompino-Marschall, B. (1989). On the psychoacoustic nature of the P-centre phenomenon. *Journal of Phonetics* 17, 175–192.

Pompino-Marschall (1991). The syllable as a prosodic unit and the so-called P-centre effect. *Vorschungsberichte des Instituts fur Phonetik und sprachliche Kommunikation der Universität Munchen* 29, 65–123.

Prince, A. (1983). Relating to the grid. *Linguistic Inquiry* 14, 9–100.

Prince, A. & P. Smolensky (1993). *Optimality theory: Constraint interaction in generative grammar.* Ms, Rutgers University and University of Colorado, Boulder.

Quené, H. (1989). The influence of acoustic-phonetic word boundary markers on perceived word segmentation in Dutch. PhD dissertation, Utrecht University.

Remijsen, B. (1996). De duurprosodische markering van de intonatiefrase in het Nederlands. [The durational prosodic marking of the intonation phrase in Dutch]. MA Thesis, Leiden University.

Revithiadou, A. (forthcoming). On the prosody/morphology interface: An investigation on head driven prominence and lexical accents. To appear as LOT/HIL PhD Dissertation. Holland Academic Graphics, The Hague.

Rietveld, A.C.M. & V.J. van Heuven (1997). *Algemene fonetiek* [General Phonetics]. Coutinho, Bussum.

Robinson, L.W. (1975). Some phonological rules of Iowa-Oto (Siouan). *Proceedings of the Mid-America Linguistics Conference*, 439–448.

Rooij, J.J. de, (1979). Speech punctuation: An acoustic and perceptual study of some aspects of speech prosody in Dutch. PhD Dissertation, Utrecht University.

Rubach, J. (to appear). The syllable in phonological analysis. To appear in: H. van der Hulst & N. Ritter (eds.) *The syllable: Views and facts*. Oxford University Press, Oxford.

Scott, S.K. (1993). P-centres in speech: An acoustic analysis. PhD Dissertation, University College London.

Selkirk, E. (1978). *On prosodic structure and its relation to syntactic structure*. Paper presented at the Conference on Mental Representation of Phonology.

Sluijter, A.M.C. (1995). Phonetic correlates of stress and accent. HIL PhD Dissertations 15. Holland Academic Graphics, The Hague.

Sluijter, A.M.C. & V.J. van Heuven (1996). Spectral balance as an acoustic correlate for linguistic stress. *JASA* 100, 2471–2485.

Sluijter, A.M.C., V.J. van Heuven & J.J. Pacilly (1997). Spectral balance as a cue in the perception of linguistic stress. *JASA* 101, 503–513.

Smith, N. (1997). Shrinking and hopping vowels in northern Cape York: Minimally different systems. In: F. Hinskens, R. van Hout & W.L. Wetzels (eds.). *Current issues in linguistic theory* 146. Variation, change and phonological theory, 267–302. John Benjamins, Amsterdam.

Stott, L.H. (1935). Time order errors in the discrimination of short tonal stimuli. *American Journal of Psychology* 18, 741–766.

Strehlow, T.H.G. (1942). Arandic phonetics *Oceania* XII, 255–302.

Sweet, H. (1902). *A primer of phonetics.* At the clarendon press, Oxford.

Takahashi, T. (1994). Constraint interaction in Aranda stress. *University College London Working Papers in Linguistics* 6, 479–508.

Taplin, G. (1879). *The folklore, manners, customs and languages of the South Australian aborigines.* Government Printer, Adelaide.

Terken, J.M.B. (1984). The distribution of pitch accents in instructions as a function of discourse structure. *Language and Speech* 27, 269–289.

Terken, J.M.B. (1991). Fundamental frequency and perceived prominence of accented syllables. *JASA* 89, 1768–1776.

Tindale, N.B. (1938). word lists (Walangama). Ms, AIATSIS, Canberra.

Tindale, N.B. (1974). *Aboriginal tribes of Australia*. University of California Press, Berkeley.

Vennemann, T. (1972). On the theory of syllabic phonology. *LB* 18: 1–18.

Vijver, R. van de (1997). The iambic issue. HIL PhD Dissertations 37. Holland Academic Graphics, The Hague.

Visch, E.A.M. (1996). The complexity of vowel lengthening in Carib: The role of closed syllables. In: M. Nespor & N. Smith (eds.) *Dam Phonology: HIL Phonology papers II*, 225–251. HIL Publications 3. Holland Academic Graphics, The Hague.

Vogten, L. (1984). Analyse, zuinige codering en resynthese van spraakgeluid. [Analysis, economical encoding, and resynthesis of speech sounds] PhD Dissertation. Eindhoven Technical University.

Vollmer, K. (1997). Koartikulation und glottale Transparenz [coarticulation and glottal transparency]. *Phonetik AIMS* 3:5. Universität Stuttgart.

Waals, J. (1996). On the temporal structure of the Dutch word onset. Ms, University of Utrecht.

Waals, J. (forthcoming). Segment durations and the syllable: An overview of the results of three production-experiments. To appear in: T. Cambier-Langeveld, A. Lipták & M. Redford (eds.) *ConSole VI Proceedings*. SOLE, Leiden.

Wever, E.G. (1949). *Theory of hearing*. Wiley & Sons, New York.

Woodrow, H. (1951). Time perception. In: S. Stevens (ed.) *Handbook of experimental phonology*. Wiley & Sons, New York.

Yallop, C. (1977). *Alyawarra: An aboriginal language of Central Australia*. AIATSIS, Canberra.

Zec, D. (1988). Sonority constraints on prosodic structure. PhD Dissertation, Stanford University.

Appendix A Mean duration values for the experiments in chapter 2

***Table AI**: mean Onset and Coda cluster durations (standard deviations between brackets) from the pilot experiment.*

	Onset		**Coda**		**Coda**	
Cluster size			Short vowel		Long vowel	
0	81.5	(15.2)			0.0	(0)
1	135.0	(8.5)	97.2	(20.3)	102.9	(14.2)
2	175.0	(9.8)	168.0	(29.4)	153.5	(26.8)
3	226.5	(29.4)	195.7	(26.3)	182.7	(19.1)

***Table AII**: mean Ascent and Descent durations (standard deviations between brackets) from the pilot experiment.*

	Ascent		**Descent**		**Descent**	
Cluster size			Short vowel		Long vowel	
0	114.1	(13.0)			0.0	(0)
1	169.4	(9.2)	165.6	(15.5)	226.0	(22.5)
2	210.4	(14.6)	220.9	(22.7)	262.8	(11.3)
3	260.4	(33.1)	253.3	(19.4)	299.4	(14.4)

***Table AIII**: mean Onset and Coda cluster durations (standard deviations between brackets) from the control experiment.*

	Onset	**Onset**	**Coda**	**Coda**
Cluster size	Short vowel	Long vowel	Short vowel	Long vowel
0	83.5 (15.0)	76.0 (14.0)		0.0 (0)
1	165.5 (17.9)	158.1 (21.2)	132.4 (35.8)	124.3 (48.1)
2	201.0 (30.3)	190.6 (25.4)	196.6 (31.7)	195.3 (66.7)
3	255.2 (42.7)	269.8 (20.0)	228.5 (35.0)	251.4 (56.8)

***Table AIV**: mean Onset and Coda cluster durations (standard deviations between brackets) obtained from the Hofhuis (1993) data.*

Cluster size	Onset		Coda	
0	46.4	(17.4)		
1	72.5	(14.8)	52.0	(19.2)
2	168.5	(18.0)	132.6	(15.7)
3	181.6	(15.6)	157.2	(30.4)

Appendix B Mean durations, percentage scores and figures for the experiments in chapter 3

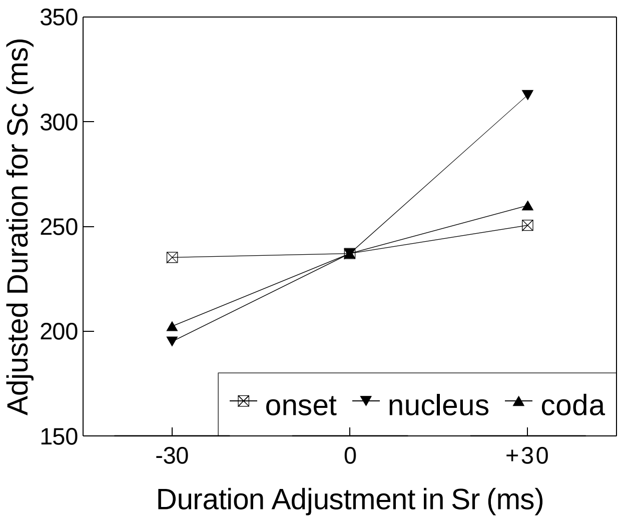


Figure 1: mean adjusted durations of Sc by duration manipulation and stimulus type (see table II)

Table BI*: mean adjusted durations of the noise burst as a function of the temporal structure of the reference syllable (figure 2/sas).*

Duration change	Manipulated Constituent Onset	Nucleus	Coda	Reference Duration
-30	159	131	149	270
none(base)	167	167	167	300
+30	179	213	173	330

Table BII*: mean adjusted durations of the noise burst as a function of the temporal structure of the reference syllable (figure 2/mam).*

Duration change	Manipulated Constituent Onset	Nucleus	Coda	Reference Duration
-30	189	171	180	270
none (base)	198	198	198	300
+30	218	242	222	330

***Table BIII**: percentages of "C longer" judgements for onset, nucleus and coda by duration manipulation (figure 4).*

Duration change	Manipulated Constituent Onset	Nucleus	Coda
-60 ms	20	8	12
-40 ms	29	9	22
-20 ms	37	27	32
base	52	52	52
20 ms	63	77	68
40 ms	69	89	80
60 ms	81	95	88

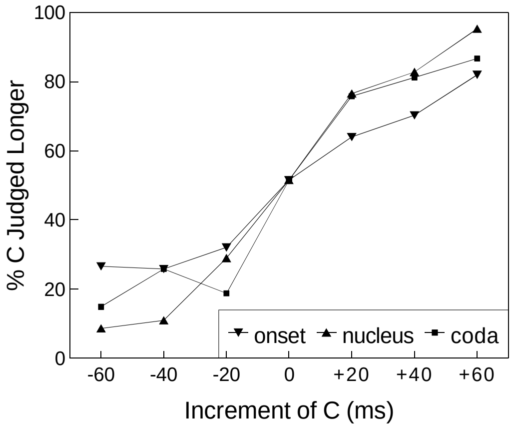


Figure 2: overall percentages of "longer" judgements for C by duration manipulation and subsyllabic constituent (base: 240 ms).

***Table BIV**: percentages of "C longer" judgements for onset, nucleus and coda by duration manipulation (base 240 ms).*

Duration change	Manipulated Constituent Onset	Nucleus	Coda
-60 ms	27	9	15
-40 ms	26	11	26
-20 ms	32	29	19
base	52	52	52
20 ms	64	76	76
40 ms	70	83	81
60 ms	82	95	88

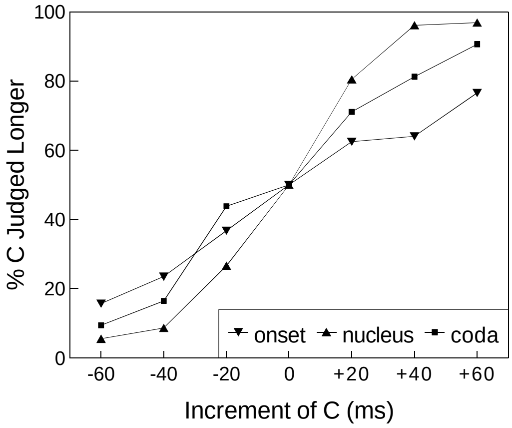


Figure 3: overall percentages of "longer" judgements for C by duration manipulation and subsyllabic constituent (base: 300 ms).

***Table BV**: percentages of "C longer" judgements for onset, nucleus and coda by duration manipulation (base 300 ms).*

Duration change	Manipulated Constituent Onset	Nucleus	Coda
-60 ms	16	5	9
-40 ms	23	9	16
-20 ms	37	27	44
base	50	50	50
20 ms	62	80	71
40 ms	64	96	81
60 ms	77	97	91

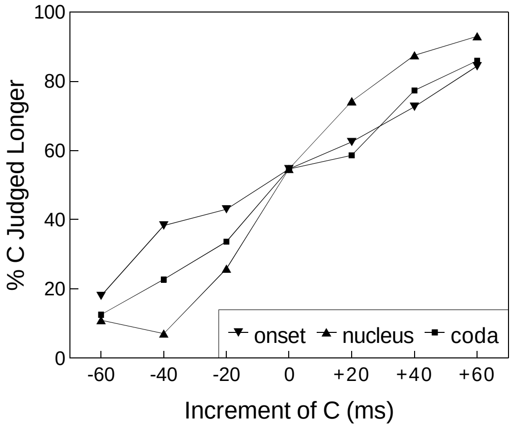


Figure 4: overall percentages of "longer" judgements for C by duration manipulation and subsyllabic constituent (base: 360 ms).

Table BVI*: percentages of "C longer" judgements for onset, nucleus and coda by duration manipulation (base 360 ms).*

Duration change	Manipulated Constituent Onset	Nucleus	Coda
-60 ms	18	11	13
-40 ms	38	7	23
-20 ms	43	26	34
base	55	55	55
20 ms	63	74	59
40 ms	73	88	77
60 ms	84	93	86

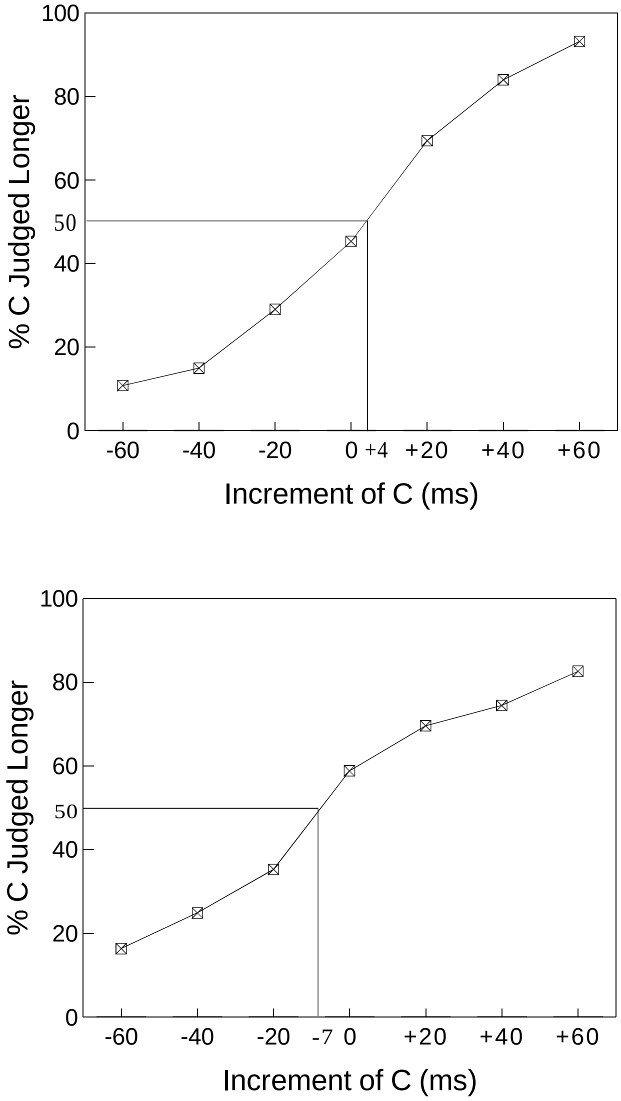


Figure 6: overall percentages of "longer" judgements for C by duration manipulation for the orders base-C (top panel) and C-base (bottom panel).

Appendix C Figures and tables for chapter 4

Schematised intensity envelopes for the stimuli in the intensity experiment, and mean durations and percentage scores for the P-centre experiment and the artificial syllable experiment from chapter 4.

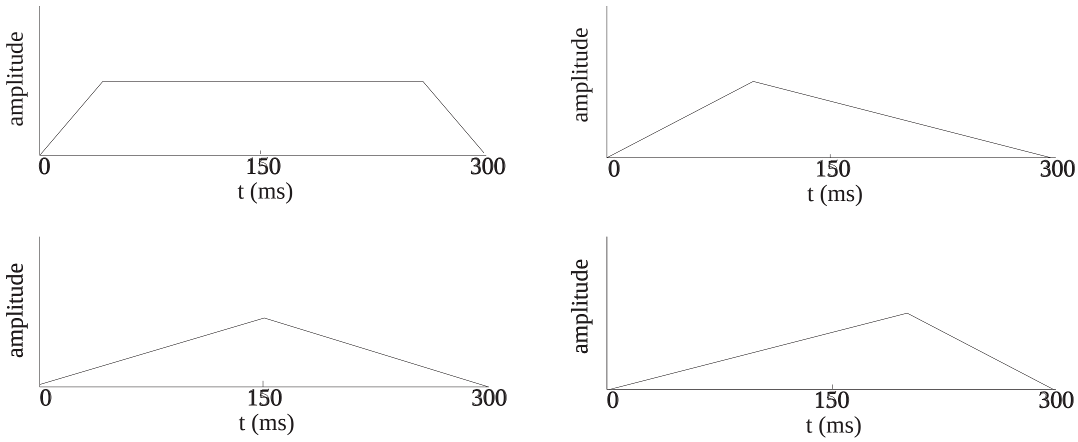


Figure 1: schematic intensity envelopes for Sr-f (top left), Sr-p (bottom left), and Sc's with peak at 100 ms (top right) and 200 ms (bottom right), see chapter 4, table I.

Table CI*: P-centre location (ms) for speech (S) and noise (N) by intensity peak position(Pk) (figure 5).*

Pk	40	60	80	100	120	140	160	180	200	220	240	260
S	92	85	94	65	97	126	76	65	103	130	192	131
N	72	87	88	89	130	115	141	116	173	150	205	180

***Table CII**: percentages of "C longer" judgements for artificial onset, nucleus and coda by duration manipulation (figure 7).*

Duration change	Onset	Nucleus	Coda
-60 ms	6	9	4
-40 ms	5	9	7
-20 ms	18	24	29
base	47	47	47
20 ms	66	65	78
40 ms	85	92	89
60 ms	93	96	95

Appendix D Australian Aboriginal languages by stress type

This is a list of languages categorised by stress type and annotated with the main descriptive source and some codes denoting special characteristics of the stress pattern in that language.

This list does not contain all the Aboriginal languages that were spoken in Australia. Descriptions of some languages were too meagre to conclude anything about the stressing, and other languages were lost before anything was written down. Some of the entries have no descriptive source. These must be handled with caution. The stress pattern in these languages was deduced from old scanty material found in the Australian Institute for Aboriginal and Torres Strait Islander Studies (AIATSIS) library and at the Australian National University (ANU). It may be that some of them have been mistakenly placed in a certain category, due to missing crucial information. Yet, they were included for completeness sake. The T-numbers behind the language names are those given to the languages by Norman Tindale in his 1938 survey (d of Tx, denotes dialect of x).

Explanation of codes:

C: stress is claimed to be contrastive
E: the descriptive source notes many exceptions to the general rule
Ed: 1 secondary stress is located at the edge opposite of main stress, or footing for secondary stress starts at that edge
ID: the language underwent Initial Dropping
M: morphological considerations play a role in stressing
NMS: No Main Stress, all the stresses in the word are equally strong
QS: marginal quantity-sensitivity (only indicated for languages that are QI otherwise)
NS: source mentions there is no sec stress
US: unclear whether there is secondary stress
ES: exceptional secondary stresses
RS: regular secondary stresses

Stress categories

I Main stress on the initial syllable, secondary stress on alternates. (allowance or disallowance of secondary stress on final syllables is not indicated)

Anguthimri (T156) (NMS):
Crowley, Terry (1981). The Mpakwithi dialect of Anguthimri. In: Bob Dixon & Barry Blake (eds.) *Handbook of Australian languages* 2, 146–194. John Benjamins, Amsterdam.
Arabana-Wangkangurru (T49) (M):
Hercus, Luise (1994). *A grammar of the Arabana-Wangkangurru language: Lake Eyre Basin, South Australia*. Pacific Linguistics, Series C-128. Canberra.
Bardi (T84):
Metcalfe, C.D. (1971). *A tentative phonetic statement of the Bardi aboriginal language*. Papers on the languages of the Australian aborigines. Australian Aboriginal Studies 38, 82–92. AIATSIS, Canberra.
Bidjara/Gungubula (T217) (M):
Breen, Gavan (1973). *Bidyara and Gungabula: Grammar and vocabulary*. Linguistic Communications 8. Monash University, Melbourne.
Biri (T213):
Terril, Angela (1993). Biri: A salvage study of a Queensland language. BA Thesis, ANU.
Bularnu (T246) (QS):
Breen, Gavan (n.d.). Bularna grammar and phonology. Ms, AIATSIS, Canberra.
Bunuba (T88):
Rumsey, Alan (n.d.). Brief tentative description of Bunaba. Ms, ANU.
Burarra (T101) (M):
Glasgow, Kathleen (1981). Burarra. In: Bruce Waters (ed.), *AAB SIL Working Papers*. Series A-5, 63–90.
Dalabon (T129) (QS: glottal stop in coda):
Capell, Arthur (1962). *Some linguistic types in Australia*. Oceania Linguistic Monographs 7. University of Sydney.
Dhuwal/Dhuwala (T251):
Amery, Robert M. (1985). A new diglossia: Contemporary speech varieties at Yirrkala in North-East Arnhem Land. MA Thesis, ANU.

Diyari (T47) (M):
Austin, Peter (1981). *A grammar of Diyari: South Australia*. Cambridge Studies in Linguistics 32. CUP, Cambridge.
Djadjala (d of Wergaia) (QS):
Hercus, Luise (1986). *Victorian languages: A late survey*. Pacific Linguistics, Series B-77. Canberra.
Dyirbal (T204):
Dixon, Bob (1972). *The Dyirbal language of Northern Queensland*. Cambridge Studies in Linguistics 9. CUP, Cambridge.
Juat:
Douglas W.H. (1968). *The aboriginal language of South-West Australia*. Australian Aboriginal Studies 14. Linguistics Series 4. AIATSIS, Canberra.
Gaanay (T14) (QS):
Hercus, Luise (1986). *Victorian languages: A late survey*. Pacific Linguistics, Series B-77. Canberra.
Garawa (T249) (Ed):
Furby, Christine (1974). Garawa phonology. Pacific Linguistics, Series A-7, 1–11. Canberra.
Gooniyandi (T89):
McGregor, W. (1990). *A functional grammar of Gooniyandi*. John Benjamins, Amsterdam.
Gugada (d of T100):
Platt, John T. (1972). *An outline grammar of the Gugada dialect: South Australia*. Australian Aboriginal Studies 48. Linguistic Series 20. AIATSIS, Canberra.
Gunin/Kwini:
McGregor, William B. (1993). *Gunin/Kwini*. Languages of the world materials 11. Lincom Europa, München.
Guṇdidj (T4) (QS):
Hercus, Luise (1986). *Victorian languages: A late survey*. Pacific Linguistics, Series B-77. Canberra.
Gunwiggu (Kunwinkju) (T125) (Ed):
Oates, Lynette F. (1964). *A tentative description of the Gunwiggu language (of Western Arnhem Land)*. Oceania Linguistic Monographs 10, University of Sydney. (This language may also belong to the penultimate stress category.)
Gupapuyngu (d of T251):
Elwell, Vanessa M.R. (1979). English as a second language in aboriginal Australia: A case study of Milingimbi. MA Thesis, ANU.

Guwa (T233):

Breen, Gavan (1971). Aboriginal languages of West Queensland, appendix. Ms, AIATSIS, Canberra.

Kaurna (T39):

Simpson, Jane p.c.

Kuku-Yalanji (T201) (Ed):

Oates, William J. & Lynette F. Oates (1964). Gugu-Yalanji linguistic and anthropological data. *Gugu-Yalanji and Wik-Mungan language studies*, 1–17. AIATSIS, Canberra.

Kurtjar (T180):

Black, Paul (1977). Kurtjar - English dictionary. Ms, ANU.

Kuthant (T179):

Black, Paul (1975). Kurtjar and Kuthant: interim report AIATSIS, Canberra.

Kuuk Thayorre (T173):

Hall, Allen H. (1968). A depth-study of the Thayorre language of the Edward River tribe, Cape York peninsula. MA thesis, University of Queensland.

Manbara (T207):

Sutton, Peter (n.d.). Ms on speech from palm island, ANU.

Mantjiltjara (d of T100):

Marsh, James (1969). Mantjiltjara phonology. *Oceanic Linguistics* 8:2, 131–152.

Maranungku:

Tryon, Darrell T. (1970). *An introduction to Maranungku*. Pacific Linguistics, Series B-14. Canberra.

Margany:

Breen, Gavan (1981). Margany. In: Bob Dixon & Barry Blake (eds.) *Handbook of Australian Languages* 2, 275–393. John Benjamins, Amsterdam.

Marrgu (T107) (E):

No reference.

Marrithiyel (T504):

Green, Ian (1989). Marrithiyel: A language of the Daly River region of Australia's Northern Territory. PhD Dissertation, ANU.

Martuthunira (T75) (QS) (NMS) (M):

Dench, Alan C. (1987). Martuthunira: A language of the Pilbara region of Western Australia. PhD Dissertation, ANU.

Mbara (T182,3) (ID):

Breen, Gavan (1979). Mbara/Yanga vocabulary/ Mbara phonology/ Mbara grammar. Ms, AIATSIS.

Sutton, Peter (n.d.). Yanga-Mbara. Ms, AIATSIS.

Miriwung (T310) (E):
Kofod, Frances M. (1978). The Miriwung Language: A phonological and morphological study. MA Thesis, University of New England.
Murrinh-patha (T510) (E) (NMS):
Walsh, Michael J. (1976). The Murinypata language of North-West Australia. PhD Dissertation, ANU.
Nakkara (T103) (Ed):
Eather, Bronwyn (1990). A grammar of Nakkara. PhD Dissertation, ANU.
Ngaanyatjara (d of T100):
See, R. (1965). Comparison of some Australian languages. PhD Dissertation, UCLA.
Ngadjuri (T42):
Berndt & Vogelsang (1941). Comparative vocabularies of the Ngadjuri and Dieri tribes, South Australia. *TRSSA* 65:1, 1–18.
Ngaliwurru (T308):
Bolt, J., W. Hoddinott & F. Kofod (1971). *An elementary grammar of the Nungali language of the Northern Territory*. Mimeo, Armadale.
Ngardi (T93):
Menning, Kathy & David Nash (1981). *Sourcebook for Central Australian languages*. Institute for Aboriginal Development, Alice Springs.
Ngawun (Mayi) (T239) (E) (M):
Breen, Gavan (1981). *The Mayi languages of the Queensland Gulf country*. AIATSIS, Canberra.
Nhanta (T56) (ID) (M):
Blevins, Juliette & Doug Marmion (1994). Nhanta historical phonology. *Australian Journal of Linguistics* 14:2, 193–216.
Nungali (T306):
Bolt, J., W. Hoddinott & F. Kofod (1971). *An elementary grammar of the Nungali language of the Northern Territory*. Mimeo, Armadale.
Nyikina (T83) (E) (QS):
Stokes, Bronwyn (1982). A description of Nyigina: A language of the West Kimberley, western Australia. PhD Dissertation, ANU.
Nyungar (T53) (Ed):
Douglas, Wilfrid H. (1976). *The aboriginal languages of the south-west of Australia*. Research and Regional Studies 9. AIATSIS, Canberra.
Paakantji (Baagandji) (T38) (E):
Hercus, Luise (1982). *The Bāgandji language*. Pacific Linguistics, Series B-67. Canberra.
Palyku (T78):
O'Grady, Geoffrey N. (1957). Balygu. Ms, AIATSIS, Canberra.

Partimaya (T59):
Dunn, Leone (1988). Badimaya: A western Australian language. Pacific Linguistics, Series A-71, 19–149. Canberra.
Pintupi (d of T100):
Hansen, K.C. & L.E. Hansen (1969). Pintupi phonology. *Oceanic Linguistics* 8, 153–170.
Pitjantjara (d of T100):
Glass, A.D. & D. Hackett (1970). *Pitjantjara grammar*. Australian Aboriginal Studies 34. Linguistics Series 13. AIATSIS, Canberra.
Pitta-Pitta (T235) (M):
Blake, Barry & Ian Green (1971). *The Pitta Pitta dialects*. Linguistic Communications 4. Monash University, Melbourne.
Blake, Barry (1979). Pitta Pitta. In: Bob Dixon & Barry Blake (eds.) *Handbook of Australian Languages* 1, 182–242. John Benjamins, Amsterdam.
Rembarrnga (T126) (NMS):
McKay, Graham R. (1975). Rembarnga: A language of central Arnhem Land. PhD Dissertation, ANU.
Waalubal (d of T28) (QS):
Crowley, Terry (1978). *The Middle Clarence dialects of Bandjalang*. AIATSIS, Canberra.
Waanji (T248) (Ed):
Osborne, C. (1966). A tentative description of the Waanji language. Ms, ANU.
Walmatjari (T99):
Hundson, Joyce (1978). *The core of Walmatjari grammar*. AIATSIS, Canberra.
Wambaya (T303):
Campbell, S.J. (1977). An outline grammar of Wambaya. Ms, ANU.
Wangkumara (T229):
McDonald, Maryalyce & Stephen A. Wurm (1979). *Basic materials in Wankumara (Galali): Grammar, sentences and vocabulary*. Pacific Linguistics, Series B-65. Canberra.
Warlpiri (T92) (M):
Nash, David G. (1980). Topics in Warlpiri grammar. PhD Dissertation, MIT.
Warumungu (T95):
Evans, Nicholas (1982). *A learner's guide to Warumungu*. Institute of Aboriginal Development, Alice Springs.
Warray (T122) (Ed) (M):
Harvey, Mark (1987). The Waray language from Adelaide River. MA Thesis, ANU.

Watjarri (T57) (Ed):
Douglas, Wilfrid H. (1981). Watjarri. In: Bob Dixon & Barry Blake (eds.) *Handbook of Australian Languages* 2, 197–272. John Benjamins, Amsterdam.
Wembawemba (T5) (E):
Hercus, Luise (1986). *Victorian languages: A late survey*. Pacific Linguistics, Series B-77. Canberra.
Western Torres Strait (Kala Lagaw Ya) (T150):
Kennedy, R.J. (1981). Kalaw Lagaw Ya. In: Bruce Waters (ed.), *AAB SIL Working Papers*. Series A-5, 103–138.
Wik-Mungknh (T169):
Sayers, Barbara J. (1974). Interpretation of stress and pitch in Wik-Munkan grammar and phonology. MS, SIL.
Wiradhurri (T32):
No reference.
Wulna (T118) (E):
No reference.
Wuywurrung (T8):
Hercus, Luise (1986). *Victorian languages: A late survey*. Pacific Linguistics, Series B-77. Canberra.
Yabala-yabala (T10):
Yallop, Colin (1975). *Narinjari: An outline grammar of the language studied by George Taplin with Taplin's notes and comparative tables*. Oceanic Linguistic Monographs 17. University of Sydney.
Yanda (T234):
Breen, Gavan (1971). Aboriginal languages of West Queensland, appendix. Ms, AIATSIS.
Yandruwanhdha (T46):
No reference.
Yankunytjatjara (d of T100) (M):
Goddard, Cliff (1983). A semantically-oriented grammar of the Yankunytjatjara dialect of the Western Desert language. PhD Dissertation, ANU.
Yaralde (T1) (E):
McDonald, Maryalyce (1977). A study of the phonetics and phonology of Yaraldi and associated dialects. MA Thesis, ANU.
Yir-Yoront (T174):
Alpher, Barry (1991). *Yir Yoront lexicon. Sketch and dictionary of an Australian language*. Trends in Linguistics documentation 6. Mouton, Berlin.
Yithayitha (T35) (E):
No reference.

II Initial main stress, no info on sec stress:

Adjnyamadhanha (T45) (ID):
Tunbridge, Dorothy (1991). *The story of the Flinders Range mammals*. Kangaroo Press, Kenthurst.
Badjiri (T227):
No reference. (Very little information was found.)
Djamindjung (T307):
Cleverly, John R. (1968). A preliminary study of the phonology and grammar of Djamindjung. MA Thesis, University of New England.
Djinang/Djinba (T255) (US):
Waters, Bruce E. (1984). Djinang and Djinba: A grammatical and historical perspective. MA Thesis, ANU.
Gaagudju (T114):
No reference.
Gungarakanj (T121):
No reference.
Kalkatungu (T238) (US):
Blake, Barry (1979). *A Kalkatungu grammar*. Pacific Linguistics, Series B-57. Canberra.
Karlamay (T63):
No reference.
Kayardild (T242):
Keen, Sandra (1969). Ms, AIATSIS, Canberra.
Ngarla:
No reference.
Ngarluma (T76):
Inferred from: O'Grady, Geoffrey N. (1957). Fieldnotes. Ms, AIATSIS, Canberra.
Nhuwala (T74):
Inferred from: O'Grady, Geoffrey N. (1958). Fieldnotes. Ms, AIATSIS, Canberra.
Nukunu (T40):
Hercus, Luise (1992). A Nukunu dictionary. AIATSIS internal publication.
Nyamal (T79):
Inferred from: word lists Tindale (1953), Geytenbeek, Ms, AIATSIS, Canberra.
Nyangumarda (T80) (NS):
Hoard, James E. & Geoffrey N. O'Grady (1976). Nyangumarda phonology, a preliminary report. In: Bob Dixon (ed.) *Grammatical Categories in Australian Languages*, 51–77. AIATSIS, Canberra.

Olgolo (ID):
Dixon, Bob (1982). Olgolo syllable structure and what they are doing about it. In: Bob Dixon (ed.) *Where have all the adjectives gone?*, 207–210. Mouton, Amsterdam.
Patjtjamalh (T508) (M):
Ford, Lysbeth J. (1990). The phonology and morphology of Bachamal (Wogait). MA Thesis, ANU.
Payungu (T70):
Austin, Peter (1992). *A dictionary of Payungu*. La Trobe university, Melbourne.
Ritharngu (T253) (NS):
Heath, Jeffrey (1980). *Basic materials in Ritharngu grammar: Texts and dictionary*. Pacific Linguistics, Series B-62. Canberra.
Thalantji (T72):
Austin, Peter (1992). *A dictionary of Thalanyji: Western Australia*. La Trobe University, Melbourne.
Tharrkari (T65):
Klokeid, Terry J. (1969). Thargari phonology and morphology. Pacific Linguistics, Series B-12. Canberra.
Tjiwarli (T67):
Austin, Peter (1992). *A dictionary of Jiwarli*. La Trobe university, Melbourne.
Umpila (T165):
Harris, Barbara D. & Geoffrey N. O'Grady (1976). An analysis of the progressive morpheme in Umpila verbs. In: Peter Sutton (ed.) *The Languages of Cape York*, 165–213. AIATSIS, Canberra.
Warriyanka (T66):
Austin, Peter (1992). *A dictionary of Warriyangga*. La Trobe university, Melbourne.
Wik-Me'enh (T168):
Patz, Elizabeth (1977). A sketch grammar of Wik Ep. Ms, ANU.
Wuthati (T154):
Seligmann, C.G. & G. Pimm (1907). Vocabulary of the Otati language spoken at Cape Grenville. *Cambridge Anthropological Expedition Reports*, Vol 3, 277–280.
Yota-Yota (T9) (E) (ES):
Hercus, Luise (1986). *Victorian languages: A late survey*. Pacific Linguistics, Series B-77. Canberra.
Yulbaridja (d of T100):
Burridge, Kate (n.d.) A sketch grammar of Yulbaridja. Ms, AIATSIS, Canberra.

IIa Initial stress with very frequent stress on second, no regularities discovered:

Darkinjung (T22):
No reference. (Very little information was found.)
Dharuk (T21):
No reference. (Very little information was found.)
Gurrgoni (T102) (QS):
Elwell, Vanessa (1977). A preliminary analysis of Gungurugoni: A language of Northern Central Arnhem Land. Ms, ANU.
Kija (T309) (C) (RS):
Taylor, Peter & John Taylor (1971). *Kitja*. Papers on the languages of the Australian aborigines. Australian Aboriginal Studies 38, 100–109. AIATSIS, Canberra.
Kok Narr (T177):
Breen, Gavan (1976). An introduction to Gog-Nar. In: Peter Sutton (ed.) *The Languages of Cape York*, 243–259. AIATSIS, Canberra.
Kok Thaw (T176) (C):
Sommer, B.A. (1971-74) Field Books. Ms, AIATSIS, Canberra.
Koko Bera (T175) (C):
No reference. (Stress on third syllables is observed, in these cases a secondary stress occurs on the initial syllable.)
Ndjebbana (T104) (C) (E):
McKay, G.R. (1979). Djeebbana phonemic statement. Ms, dept of education. Maningrida, Northern Territory.

III Stress is on the first or second syllable, depending on the number of syllables in the word:

Malakmalak:
Birk, D.B.W. (1976). *The Malakmalak language: Daly River (Western Arnhem Land)*. Pacific Linguistics, Series B-45. Canberra.
Ngankikurrunggurr (Ngan'gitjemerri) (T509):
Hoddinott, W.G. & F.M. Kofod (1988). *The Ngankikurungkurr language* (Daly River area, Northern Territory). Pacific Linguistics, Series D-77. Canberra.
Nyawaygi (T206) (QS):
Dixon, Bob (1983). Nyawaygi. In: Bob Dixon & Barry Blake (eds.) *Handbook of Australian Languages* 3, 430–525. John Benjamins, Amsterdam.

Warrgamay (T205) (QS):
Dixon, Bob (1981). Wargamay. In: Bob Dixon & Barry Blake (eds.) *Handbook of Australian Languages* 2, 1–144. John Benjamins, Amsterdam.

IV Stress is on the penultimate syllable:

Alawa (T139) (E):
Sharpe, Margaret C. (1972). *Alawa phonology and grammar*. AIATSIS, Canberra.
Jingulu (T301):
Chadwick, Neil (1975). *A descriptive study of the Djingili language*. Research and Regional Studies 2. AIATSIS, Canberra.
Maung (T105) (E):
Capell, Arthur & H.E. Hinch (1970) *Maung grammar: Texts and vocabulary*. Mouton, The Hague.
Ngalakan (T128) (E):
Merlan, Francesca (1983). *Ngalakan grammar, Texts and vocabulary*. Pacific Linguistics, Series B-89. Canberra.
Ngalkbun (E) (Ed):
Sandefur, John & David Nangan:golod Jenhan (1977). A tentative description of the phonemes of the Ngalkbun language. *Workpapers of SIL AAB.* Series A-1, 57–96.
Ngarndji:
Chadwick, Neil (1971). Ngarndji word list and phonological key. In: *Papers on the Languages of Australian Aborigines*, 34–45. Australian Aboriginal Studies 38. AIATSIS, Canberra.
Nunggubuyu (T135) (QS) (Ed):
Hore, Michael (1981). Syllable length and stress in Nunggubuyu. In: Bruce Waters (ed.), *AAB SIL Work Papers*. Series A-5, 1–62.
Oykangand (T186) (E = final stress?)
Sommer, B.A. (1969). *Kunjen phonology: Synchronic and diachronic*. Pacific Linguistics, Series B-11. Canberra.
Tiwi (T123) (Ed):
Lee, Jennifer (1987). *Tiwi today: A study of language change in a contact situation*. Pacific Linguistics, Series C-96. Canberra.
Osbourne, C.R. (1974). *The Tiwi language*. Australian Aboriginal Studies 55. Linguistic Series 21. AIATSIS, Canberra.
Umbugarla (T116) (E) (Ed):
Davies, Jennifer (1989). Umbugarla: A sketch grammar. Honours Thesis, ANU.

Wardaman (T132) (NMS) (M):
Merlan, Francesca (1994). *A grammar of Wardaman: A language of the Nortern Territory of Australia*. Mouton, Berlin-NY.
Yanyuwa (T250) (NMS):
Kirton, J.F. (1977). *Anyula phonology*. Pacific Linguistics, Series A-10. Canberra.

V Stress is on the second syllable, three related cases, and one neighbouring language:

Gabi-Gabi (T222):
Holmer, Nils M. (1983). *Linguistic survey of South-Eastern Queensland*. Pacific Linguistics, Series D-54. Canberra.
Gureng-Gureng (T221):
Holmer, Nils M. (1983). *Linguistic survey of South-Eastern Queensland*. Pacific Linguistics, Series D-54. Canberra.
Waga-Waga (T223):
Holmer, Nils M. (1983). *Linguistic survey of South-Eastern Queensland*. Pacific Linguistics, Series D-54. Canberra.
Yagara (T225):
No reference.

VI Stress is on the second syllable if it is heavy, otherwise on the first:

Anindhilyagwa (T136) (NMS) (limited to 3 syllable words, QS: vowel quality):
Stokes, Judith (1981). Anindilyakwa. In: Bruce Waters (ed.), *AAB SIL work papers*. Series A-5, 1–62.
Bandjalang (Waalubal dialect) (T28):
Crowley, Terry (1978). *The Middle Clarence dialects of Bandjalang*. AIATIS, Canberra.
Birbay (d of T24):
No reference.
Dharawal (T18):
Eades, Diana K. (1976). *The Dharawal and Dhurga languages of the New South Wales coast.* Australian Aboriginal Studies. Research and Regional Studies 8. AIATSIS, Canberra. (Described as "first long vowel or first" but long vowels beyond the second syllable are not found.)

Dhurga (T17):
Eades, Diana K. (1976). *The Dharawal and Dhurga languages of the New South Wales coast.* Australian Aboriginal Studies. Research and Regional Studies 8. AIATSIS, Canberra. ("First long vowel or first": see Dharawal entry)
Djabugay (T202) (limited to 2 and 3 syllable words, avoids final stress):
Patz, Elizabeth (1991). Djabugay. In: Bob Dixon & Barry Blake (eds.) *Handbook of Australian Languages* 4, 245–347. OUP, Oxford.
Gidabal (d of T28) (QS):
Geytenbeek, Brian & Helen Geytenbeek (1971). *Gidabal grammar and dictionary*. Australian Aboriginal Studies 43. Linguistic Series 17. AIATSIS, Canberra.
Guugu Yimidhirr (T200):
Haviland, John (1979) Guugu Yimidhirr. In: Bob Dixon & Barry Blake (eds.) *The Handbook of Australian Languages* 1, 27–180. John Benjamins, Amsterdam. (Initial *and* second if second is heavy, else initial.)
Gumbaynggirr (T26) (E):
Eades, Diana (1979). Gumbayngir. In: Bob Dixon & Barry Blake (eds.) *Handbook of Australian Languages* 1, 244–361. John Benjamins, Amsterdam.
Ngarinyin (T86) (E):
Coate, H.H.J. & Lynette Oates (1970). *A grammar of Ngarinjin: Western Australia.* Australian Aboriginal Studies 25. Linguistic Series 10. AIATSIS, Canberra.
Thawa (T15):
Eades, Diana K. (1976). *The Dharawal and Dhurga languages of the New South Wales coast.* Australian Aboriginal Studies. Research and Regional Studies 8. AIATSIS, Canberra.
Wagaya (T247) (E):
No reference. (Sensitive not only to vowel length but also vowel height. Many exceptions. This language may also belong in the "unpredictable", initial or second category.)
Walangama (T181) (NMS):
No reference.
Yaygirr (T27) (ID):
Crowley, Terry (1979). Yaygir. In: Bob Dixon & Barry Blake (eds.) *Handbook of Australian Languages* 1, 363–390. John Benjamins, Amsterdam.
Yintjiparnti (Kurama) (T76):
Wordick, F.J.F. (1982). *The Yindjibarndi language*. Pacific Linguistics, Series C-71. Canberra.

VII Stress is on the first long vowel, else on the first syllable:

Djaru (T98) (M):
Tsunoda, Tasaku (1981). *The Djaru language family of Kimberley, Western Australia*. Pacific Linguistics, Series B-78. Canberra.
Gamilaraay (Kamilaroy) (T31):
Austin, Peter (1993). A reference dictionary of Gamilaraay, northern New South Wales. Bundoora. La Trobe University, Melbourne. (Has two main stresses if two syllables have a long vowel.)
Kukatj (T178) (Ed) (QS: vowel quality, a word never consists of only light syllables):
Breen, Gavan (1992). Some problems in Kukatj phonology. *Australian Journal of Linguistics* 12:1, 1–44.
Muruwarri (T34) (Ed):
Oates, Lynette F. (1988). *The Muruwari language*. Pacific Linguistics, Series C-108. Canberra.
Ngiyampaa (T33) (M) (Ed):
Donaldson, Tasmin (1977). A description of Ngiyamba: The language of the Wanga:ybuwan people of central New South Wales. PhD Dissertation, ANU.
Panytjima (T76) (M) (Ed) (RS):
Dench, Alan C. (1981). Panjima phonology and morphology. MA Thesis, ANU.
Dench, Alan C. (1991). Panyjima. In: Bob Dixon & Barry Blake (eds.) *Handbook of Australian Languages* 4, 124–243. OUP, Oxford.
Yukulta (T243) (M) (Ed):
Keen, Sandra (1983). Yukulta. In: Bob Dixon & Barry Blake (eds.) *Handbook of Australian Languages* 3, 191–305. John Benjamins, Amsterdam.
Yuwalaraay:
Williams Corinne J. (1980). A grammar of Yuwaalaraay. Pacific Linguistics, Series B-74. Canberra. (Has two main stresses if first and second syllable both have a long vowel.)

VIII Stress is on the first postconsonantal syllable:

Agwamin (T184):
No reference.
Alyawarre (d of T90) (ID):
Yallop, Colin. (1977). *Alyawarra: An aboriginal language of Central Australia*. AIATSIS, Canberra.

Arrernte (T90) (ID):

Strehlow, Theodore. (1942). Arrandic phonetics. *Oceania* XII: 255–302.

Kaytetj (T91) (ID):

Koch, Harold (1984). The category of ‘associated motion’ in Kaytej. *Language in Central Australia* 1, 23-34.

Koch, Harold (1997). Pama-Nyungan reflexes in the Arandic languages. In: Darrel Tryon & Michael Walsh (eds.) *Boundary rider: Essays in honour of Geoffrey O'Grady*, 271-302. Pacific Linguistics, C-136. Canberra.

Kuku-mini (T189) (ID?):

Inferred from: Dixon, Bob (1964). Word list in Koko Mini recorded from Mar Grunji at Wrotham park. Ms, AIATSIS.

Kuku-thaypan (T192) (ID):

Rigsby, Bruce. (1976). Kuku-Thaypan descriptive and historical phonology. In: Peter Sutton (ed.) *The languages of Cape York*, 68–77. AIATSIS, Canberra.

Lamalama (T196) (ID):

Laycock, Donald (1970). *Three Lamalamic languages of North Queensland.* Papers in Australian Linguistics 4, 71–97. Pacific Linguistics, Series C. Canberra.

Linngtigh (ID):

Hale, Kenneth (n.d.). Linngtigh. Ms, AIATSIS, Canberra.

O'Grady, G.N., C.F. Voegelin & F.M. Voegelin (1966). Languages of the world: Indo-Pacific fascicle six. *Anthropological Linguistics* 8:2.

Luthig (T155) (ID):

No reference.

Mbabaram (T185) (ID):

Dixon, Bob (1991). Mbabaram. In: Bob Dixon & Barry Blake (eds.) *Handbook of Australian Languages* 4, 349–402. OUP, Oxford.

Ogh-undjan (T187) (ID):

Inferred from: Alpher, Barry J. (1978). Unpublished “Ogundjan” data. Ms, AIATSIS, Canberra.

Rimang-gudinhma (T195):

No reference.

Takalak (T188):

No reference.

Umbuygamu/Morroba-lama (T194) (ID):

Ogilvie, Sarah (1994). The Morrobalama (Umbuygamu) language of Cape York, Australia. MA Thesis, ANU.

IX Exceptional cases:

Djapu (Initial stress but stress is on the second if it starts with /d̪/):
Morphy, Frances (1983). Djapu, A Yolngu dialect. In: Bob Dixon & Barry Blake (eds.) *Handbook of Australian Languages* 3, 1–188. John Benjamins, Amsterdam.

Guugu Ya'u (stress the last heavy syllable, else the first light one):
Thompson, David A. (1976). A phonology of Kuuku-Yaʔu. In: Peter Sutton (ed.) *The Languages of Cape York*, 213–235. AIATSIS, Canberra.

Larrikiya (T119) (Stress is on second if heavy or on third if heavy, else on first, unbounded but limited to first 3 syllables).
Capell, Arthur (1984). The Laragia language. Pacific Linguistics, Series A-68, 55–106. Canberra.

Mangarrayi (T140) (penultimate, shift to antepenultimate or initial):
Merlan, Francesca (1989). *Mangarayi*. Routledge, New York.

Mathimathi (M) (stress seems to be on the second when it has a coronal onset):
Hercus, Luise (1986). *Victorian languages: A late survey*. Pacific Linguistics, Series B-77. Canberra.

Miriam Mir (East Torres) (T151) (pitch accent system):
Piper, Nick (1989). A sketch grammar of Meryam Mir. MA Thesis, ANU.

Ngarigo (T19) (initial or second seems determined by onset second):
Hercus, Luise (1986). *Victorian languages: A late survey*. Pacific Linguistics, Series B-77. Canberra.

Parnkalla (T43) (stress is claimed to be antepenultimate, this is doubtful since this language is not spoken in the area where such right edged stress occurs, all the surrounding languages have left edged stress).
Schürmann, Clamor W. (1844). *Parnkalla Language*. Dehane, Adelaide.

Umbindhamu (T193) (Stress is variable):
No reference.

Uradhi (T153) (stress on heavy syllables, else on the antepenult):
Crowley, Terry (1983). Uradhi. In: Bob Dixon & Barry Blake (eds.) *Handbook of Australian Languages* 3, 306–428. John Benjamins, Amsterdam.

Yidin^y (T203) (reported to be iambic with trochaic default):
Dixon, Bob (1977). *A Grammar of Yidin^y*. Cambridge Studies in Linguistics 19. CUP, Cambridge.

Samenvatting in het Nederlands

Veel van de talen die op onze wereld worden gesproken, kennen het verschijnsel klemtoon. In tegenstelling tot wat veel mensen denken is de positie van die klemtoon binnen het woord meestal niet willekeurig. Voor de meeste klemtoontalen zijn er regels op te stellen die de positie van de klemtoon voorspellen. Voor het Nederlands lukt dat in ca. 85% van de gevallen, de rest van de woorden zijn uitzonderingen. In sommige andere talen zijn er geen uitzonderingen en kan men voor alle woorden de plaats van de klemtoon voorspellen.

De regels die in de metrische fonologie worden gebruikt om de positie van de klemtoon te voorspellen, verwijzen vaak naar bepaalde lettergrepen in het woord. Het Tsjechisch, bijvoorbeeld, plaatst klemtoon altijd op de eerste lettergreep, het Pools doet dat op de voorlaatste. Ingewikkelder wordt het als de regels niet alleen de plaats, maar ook de vorm van de lettergrepen in het woord 'meewegen' in hun beslissing. Een typisch voorbeeld van zo'n klemtoonregel zou kunnen zijn: "klemtoon valt op de laatste lettergreep als die een lange klinker bevat, anders op de voorlaatste". Deze regel vinden we in het Rotuman (een Polynesische taal). Wat ook kan, en feitelijk gebeurt in het Sentani (een Papua-taal uit Nieuw Guinea), is dat de klemtoon valt op de laatste lettergreep als die afgesloten wordt met een medeklinker, en anders op de voorlaatste. Combinaties van deze twee soorten regels zijn mogelijk; bovendien zijn ze ook toepasbaar elders in het woord (bijvoorbeeld op de eerste en de tweede lettergreep). Wat echter niet mogelijk lijkt te zijn is een taal met klemtoonregels die verwijzen naar de *begin*medeklinker(s) van de lettergreep (maar zie de samenvatting van hoofdstuk 5). Met andere woorden, er zouden geen regels zijn als: "klemtoon valt op de laatste lettergreep als die met een medeklinker begint, anders op de voorlaatste".

In het vakjargon vormt de klinker (of de klinkers) van een lettergreep de *nucleus*, (de *e* in 'sterk') de slotmedeklinker(s) de *coda* (*rk* in 'sterk') en de beginmedeklinkers de *onset* (*st* in 'sterk'). De nucleus en de coda vormen samen het *rijm* (dat wat gelijk moet klinken in woorden die rijmen). Uit de bovengenoemde observaties blijkt dus dat alleen de segmenten die in het rijm zitten mee mogen doen in de klemtoonregels; ze hebben fonologisch *gewicht*. De onset doet niet mee en is *gewichtloos*.

De centrale vraag waarop ik in dit proefschrift het antwoord heb gezocht is: “waarom is de onset van de lettergreep gewichtloos?”

In hoofdstuk 2 wordt fonologisch gewicht gelijkgesteld aan fonetische duur. In een productie-experiment testen we de hypothese dat de duur van gewichtloze delen van de lettergreep wel eens relatief invariant zou kunnen zijn, of met andere woorden, dat de duur van de onset altijd min of meer gelijk blijft, hoeveel segmenten er ook in zitten, terwijl de duur van de coda (sterk) toeneemt met het aantal segmenten dat er in zit. Deze relatieve duur-invariantie zou een mogelijke verklaring kunnen vormen voor de gewichtloosheid van de onset. Als de onset niet variabel is in duur, en duur is gewicht, dan heeft het geen zin om tijdens het wegen van de lettergreep de onset mee te tellen: die verandert de duur van de lettergreep toch nooit. In het experiment laten we een aantal sprekers lettergrepen produceren met een variabel aantal segmenten in onset en coda, bijvoorbeeld: ‘sop’, ‘stop’, ‘strop’; ‘laf’, ‘lafs’ en ‘lafst’. We voorspelden dat de duur van de onset gelijk zou blijven ondanks de groei in segment-aantal, maar dat de duur van de coda zou toenemen. Na duurmeting van de opgenomen uitingen bleek dit niet het geval. De duur van onset en coda nam in gelijke mate toe met het aantal segmenten. Een relatief gelijk blijvende duur kan dus geen verklaring vormen voor de gewichtloosheid van de onset.

In hoofdstuk 3 blijven we bij de aanname dat gewicht kan worden vertaald in duur en doen we een perceptie-experiment. De duur van de onset mag dan wel variabel zijn als we die in spraak meten, maar we weten nog niet hoe we die onset-duren waarnemen. Het is zeker mogelijk dat de variatie in onset-duur die we hebben gemeten niet of nauwelijks hoorbaar is. Dat zou betekenen dat de onset perceptief invariant is, hetgeen voldoende zou zijn als verklaring voor gewichtloosheid, omdat duurverschillen die we niet kunnen horen logischerwijs ook geen rol kunnen spelen in de fonologie. We zoeken ook hier weer naar relatieve invariantie. We verwachten namelijk niet dat duurverschillen in onsets helemaal niet gehoord kunnen worden, maar dat ze in ieder geval veel slechter waarneembaar zijn dan even grote verschillen in nucleus en coda.

In een aantal perceptie-experimenten hebben we systematisch de duur van onset, nucleus en coda van de Nederlandse woorden ‘mam’ en ‘sas’ gevarieerd. De verwachting was dat de *natuurgetrouwheid* waarmee zulke veranderingen gedetecteerd konden worden (veel) hoger zou zijn voor de nucleus en coda dan voor de onset. In een aantal instelexperimenten verzochten we de luisteraars een ruis-signaal op het gehoor gelijk te maken aan de ‘mam’- en ‘sas’-stimuli met de gevarieerde duren. Als voorbeeld geven we de resultaten van het eerste experiment. Een verandering van 30 ms in de onset resulteerde hier in het feit dat de proefpersonen het vergelijkingssignaal slechts met gemiddeld 8 ms aanpasten (ze horen een verandering van 30 ms dus als een verandering van 8 ms). Aanpassingen van 30 ms in de nucleus en coda resulteerden in aanpassingen van, respectievelijk, 59 en 29 ms. Het is duidelijk dat de duuraanpassing in de onset het slechtst werd waargenomen. De duurvariatie in de coda werd dus getrouw gereproduceerd; variaties in de nucleus werden met een factor 2 overschat. De perceptieve overdrijving van de nucleus-duur werd geweten aan de hoge sonorantie van de klinkers die de nucleus vullen.

In een vervolgexperiment hebben we aangetoond dat ook de *nauwkeurigheid* waarmee onset-duurverschillen worden waargenomen laag is ten opzichte van de waarneming van duurverschillen in nucleus en coda. In dit experiment maakten we paren van stimuli die bestonden uit een onveranderlijke referentielettergreep van 300 ms en een vergelijkingslettergreep waarin de duur van onset, nucleus of coda met 20, 40 of 60 ms was verlengd of verkort. Deze paren werden aangeboden aan luisteraars die telkens de langste van de twee moesten aanwijzen. Het succes waarmee de proefpersonen die taak kunnen uitvoeren vormt een maat voor de nauwkeurigheid van duurwaarneming. Die maat is het *Just Noticeable Difference* (JND), of de kleinste duurverandering die nog betrouwbaar gehoord kan worden. Uitgedrukt als een percentage van de duur van de referentiestimulus zijn de JNDs voor nucleus, coda en onset 6,6%, 10,6% en 16,2%. Duidelijk is dat de duurverandering in onsets groot moet zijn om gehoord te worden.

De slechte waarneming van onset-duren zien we dus terug in een slechte schatting van de grootte van een verandering in die onset-duur en in de lage nauwkeurigheid waarmee duurverschillen in de onset worden waargenomen. We concluderen daarom dat de gewichtloosheid van de onset veroorzaakt wordt door de *luisteraar*, die het er nogal slecht vanaf brengt als het gaat om luisteren naar duur (de fonetische vertaling van gewicht) in onsets.

Het antwoord dat we nu hebben op onze onderzoeksvraag is maar ten dele bevredigend. We weten namelijk nog niet *waarom* duurverschillen in onsets zoveel slechter waargenomen worden dan gelijkwaardige duurverschillen in de andere delen van de lettergreep. Het antwoord op die vraag geeft ons de echte reden voor de gewichtloosheid van de onset.

In hoofdstuk 4 voeren we aanvullende (psychofysische) experimenten uit waarmee we het antwoord op de vervolgvraag proberen te vinden. We proberen een oorzaak aan te wijzen voor de slechte duurperceptie in onsets. Het zou bijvoorbeeld zo kunnen zijn dat de accurate perceptie van duur pas begint nadat een bepaalde *trigger* het mechanisme waarmee we duur waarnemen in werking zet. Het spreekt voor zich dat, als de trigger een bepaalde gebeurtenis in de lettergreep is, het gedeelte na de trigger nauwkeuriger wordt waargenomen dan het gedeelte ervoor. Dus, als een opvallend punt na de onset als trigger dient, hebben we een verklaring voor de slechte duurwaarneming in onsets. Drie kandidaten komen in aanmerking: het intensiteitsmaximum, het P-centrum (of het punt in de tijd waarop de lettergreep wordt waargenomen) en de overgang tussen de onset en de nucleus (ook wel CV-transitie of pivot genoemd). Om er achter te komen of zo’n kandidaat ook daadwerkelijk als trigger voor duurwaarneming functioneert, kijken we of het verschuiven ervan in een lettergreep van overigens vaste duur enig effect heeft. Omdat alleen de duur van dat deel van de lettergreep dat na de trigger komt nauwkeurig wordt waargenomen verwachten we dat de totale waargenomen lettergreepduur kleiner wordt naarmate de trigger verder naar rechts in de lettergreep schuift.

In het eerste experiment in hoofdstuk 4 testten we de potentie van het intensiteitsmaximum. Er werden stimuli gegenereerd waarin de locatie van het maximum over bijna de gehele lettergreep varieerde. Met deze stimuli voerden we een paarsgewijs vergelijkingsexperiment en een instelexperiment uit om te zien of de variabele intensiteitspiek effect had op de waargenomen duur van de lettergreep. Dat was niet het geval. De proefpersonen hoorden alle lettergrepen als even lang.

Met betrekking tot de tweede mogelijke trigger, het P-centrum, hoeven we nu geen nieuw experiment op te zetten. Volgens de literatuur hangt de locatie van dit P-centrum onder andere af van de positie van het intensiteitsmaximum. Dus, de stimuli die we in het intensiteitsexperiment gebruikten zouden variabele P-centra moeten hebben. We weten al dat deze stimuli allemaal een perceptief gelijke duur hebben, dus als de positie van het P-centrum inderdaad varieert, heeft ook deze variatie geen invloed op de waargenomen duur. Om deze conclusie kracht bij te zetten controleren we in het voorlaatste experiment of de P-centra in onze stimuli zich gedragen zoals de literatuur voorspelt. De resultaten laten duidelijk zien dat het P-centrum naar achter in de lettergreep beweegt, in harmonie met het intensiteitsmaximum. Omdat deze verschuivingen niet hebben geleid tot een verandering in de waargenomen lettergreepduur kan ook het P-centrum de veronderstelde trigger niet zijn.

De derde kandidaat, de CV-transitie, zou nu de trigger kunnen zijn die we zoeken, maar dit is niet direct vast te stellen. Het is namelijk onmogelijk om de CV-transitie in de lettergreep te verschuiven zonder dat de duren van de onset en de nucleus veranderen. We weten al dat het veranderen van onset- en nucleus-duur invloed heeft op de totale waargenomen duur van de lettergreep, maar we kunnen niet beslissen of dit komt door de veranderde positie van de CV-transitie of door een ander effect dat wordt veroorzaakt door de (relatief) van duur veranderde onset. Omdat we hier op een dood spoor uitkwamen richtten we onze aandacht op een ander punt, zonder de hypothese dat de pivot als duurtrigger zou kunnen werken compleet af te schrijven.

In ons laatste experiment probeerden we te bepalen of spraak en niet-spraak fundamenteel van elkaar verschillen als het gaat om duurperceptie. De verschillen in duurwaarneming die we vonden in hoofdstuk 3 zouden specifiek kunnen zijn voor spraak, en dus behoren tot het domein van de taalkunde. De voorspelling die daar onmiddellijk aan gekoppeld kan worden is dat voor niet-spraakgeluiden de perceptie van duur gelijk moet zijn voor alle samenstellende delen van een complexe stimulus (een kunstmatige lettergreep) omdat ons taalcentrum niet actief is wanneer we naar zo’n stimulus luisteren.

We fabriceerden dus een niet-spraak “lettergreep” uit een tooncomplex (zaagtandtrilling; de “nucleus”) die werd voorafgegaan en gevolgd door een stukje ruis (de “onset” en “coda”). Met deze kunstmatige lettergreep herhaalden we het paarsgewijze vergelijkingsexperiment uit hoofdstuk 3. We vonden geen verschil in duurwaarneming van “onset”, “nucleus” en “coda”. Dit was wellicht te wijten aan de beperkte opzet van ons laatste experiment, maar een voorlopige conclusie (in ieder geval tot er meer experimenten zijn uitgevoerd met niet-spraak lettergrepen) zou kunnen zijn dat de verschillen die we hebben gevonden in de duurwaarneming van de onset, nucleus en coda specifiek zijn voor spraak, en dus een abstract taalkundige basis hebben.

De conclusies van onze experimenten leiden tot een voorlopig fonologisch model van de gewichtsverdeling in een lettergreep. Het abstract taalkundige vooroordeel dat luisteraars hebben voor nucleus en coda, vertalen we naar “gewichtsbouwstenen” voor deze twee constituenten. We hebben ook aangetoond dat bepaalde fysisch waarneembare eigenschappen van de lettergreep het gewicht kunnen verhogen. We hebben voorgesteld dat alleen prosodisch relevante eigenschappen van de lettergreep (zoals lexicale toon, sonoriteit, etc.) bij kunnen dragen aan lettergreepgewicht in ons model. De bouwstenen voor nucleus en coda moeten dus worden “overgoten” met een extra laagje dat de potentie heeft om de lettergreep te verzwaren. Deze combinatie noemen we *prominentie*. Fonologische regels kunnen de prominentie van de lettergreep gebruiken om een binair onderscheid te maken tussen lichte (weinig prominentie) en zware (veel prominentie) lettergrepen. De grens tussen licht en zwaar verschilt per taal en wordt voorgesteld als een prominentiedrempel in het lettergreepmodel waarboven een lettergreep uit moet komen om zwaar te zijn. In de meeste talen ligt de drempel zodanig dat alleen de abstracte bouwstenen relevant zijn (alle stenen komen boven de drempel uit en maken de lettergreep zwaarder), maar sommige talen gebruiken de fysisch waarneembare eigenschappen van de lettergreep voor een fijnmaziger onderscheid (alleen lettergrepen waarvan het extra laagje hoog genoeg is komen boven de drempel uit).

In hoofdstuk 5 stappen we over op de fonologie. Het belang van deze stap blijkt als we de klemtoonregels die voorkomen in de talen van de wereld, nader beschouwen. In tegenstelling tot wat eerder werd beweerd blijken er toch enkele talen te bestaan waarin onsets een rol lijken te spelen in de regels voor klemtoonplaatsing. Dat is op zich wel verenigbaar met wat we gevonden hebben in het voorgaande: duurwaarneming in onsets is zwak, maar niet onmogelijk, en dus is de onset een slechte kandidaat voor gewicht, maar geen onmogelijke. Echter, de meest concrete fonologisering van het feit dat duur in onsets slecht wordt waargenomen zou er een zijn waarin onset-gewicht helemaal niet wordt toegestaan. Het behouden van een metrische theorie waarin de mogelijke klemtoonregels aan banden worden gelegd door een restrictie op het voorkomen van onset-gewicht is een doel op zich (hoe eenvoudiger de theorie, hoe aannemelijker). Daarom is het een poging waard om te bezien of de talen die de restrictie op onset-gewicht lijken te schenden, op een andere manier geanalyseerd kunnen worden, zodanig dat ze in een theorie passen die het bestaan van onset-gewicht niet erkent.

Na een algemene introductie in de metrische fonologie, die geheel in het kader staat van de talen die gesproken werden (en worden) door de Australische Aborigines, volgt een overzicht van de talen die onsets lijken te gebruiken in hun klemtoonregels. De meerderheid van die talen is Australisch. Een groep van 14 talen uit het noordoosten van Australië deelt een klemtoonregel die hoofdklemtoon plaatst op de eerste lettergreep die met een onset begint. In de praktijk is dat de eerste of de tweede lettergreep. Deze regel vinden we elders in de wereld slechts in één andere taal, nl. het Banawá (een Amazonetaal). Het lijkt er sterk op dat, in deze talen, lettergrepen met een onset zwaar zijn en lettergrepen zonder onset licht. De klemtoon in deze talen zou op de eerste zware lettergreep vallen. Maar aangezien onsets niet bij mogen dragen aan lettergreepgewicht kunnen we zo'n regel niet formuleren. Een alternatief is echter voorhanden.

Uit de groep van 14 Australische talen nemen we het West Aranda en het Alyawarra als representatieve voorbeelden. We richten onze aandacht op het feit dat in deze talen alléén de eerste lettergreep onsetloos kan zijn. Dat stuurt ons in de richting van een analyse waarin de klemtoonregel voor deze talen eerder het gevolg is van zogenaamde 'randverschijnselen' dan van een structureel verschil in gewicht tussen de eerste twee lettergrepen. Een ingrijpend proces in de historie van het Aranda en het

Alyawarra dat te maken heeft met de linker woordrand, is het wegvallen van initiële medeklinkers. We geloven dat de basis voor de uitzonderlijke klemtoonregel in de betreffende talen hier ligt. Voordat de eerste consonant wegviel hadden deze talen namelijk gewoon hoofdklemtoon op de eerste lettergreep, net als de meeste andere Australische talen. Het wegvallen van de consonant had een kettingreactie tot gevolg. Als de klinker in de eerste lettergreep kort was viel die ook weg, en als de klinker lang was werd die verkort of vervangen door /a/. Het gevolg was dat, in woorden die een korte klinker hadden, de klemtoon *verplicht* naar de tweede lettergreep schoof. In woorden die een lange klinker hadden gebeurde dit ook, waarschijnlijk omdat de eerste onsetloze lettergreep niet "volledig" genoeg was om klemtoon te dragen, of gewoon analoog aan de andere woorden. In ieder geval was het resultaat van de verschuiving een situatie waarin woorden die een korte klinker hadden, klemtoon op de eerste lettergreep kregen (die vóór klinkerverlies de tweede lettergreep was), terwijl woorden die een lange klinker hadden klemtoon op de tweede lettergreep kregen (omdat de eerste niet helemaal verdween). Dit proces is nog eens verduidelijkt in (1).

(1)	vòòr consonant deletie	na consonant deletie
a.	*p*ápapa	pápa
b.	*p*áapapa	apápa

Het lijkt nu inderdaad alsof een synchrone regel voor woordklemtoon gebruik maakt van de aanwezigheid van de onset. We zien echter duidelijk in de historie van de talen in kwestie dat hier een klemtoonregel die niet verwijst naar gewicht aannemelijker is.

In hoofdstuk 6 bespreken we een analyse (in het kader van de Optimaliteitstheorie) die geen gebruik maakt van gewicht en meer recht doet aan de historische feiten. Aangenomen wordt dat, in talen als het Aranda en het Alyawarra, een voet (een bouwsteen van twee lettergrepen die wordt gebruikt in metrische analyses) moet beginnen met een onset. In het Aranda en het Alyawarra valt de hoofdklemtoon op de eerste lettergreep in de eerste voet. In (1a)-woorden kunnen we die eerste voet probleemloos zetten op de eerste twee lettergrepen zodat klemtoon op de eerste terechtkomt. In (1b)-woorden heeft de eerste lettergreep geen onset en is dus ongeschikt om als eerste in een voet te staan. Die lettergreep wordt dus overgeslagen en de voet schuift op naar de volgende twee lettergrepen, zodat hoofdklemtoon op de tweede lettergreep terechtkomt. Onze claim is dat deze analyse geldt voor alle Australische talen in de eerder genoemde groep van 14. Deze talen hebben namelijk zonder uitzondering allemaal het consonantdeletieproces in hun historie.

In enkele andere talen die worden besproken in hoofdstuk 5, is niet de *aanwezigheid* maar de *identiteit* van de onset de bepalende factor in de klemtoonregel. In met name het Pirahã (Amazone) en het Puluwat (Micronesië) lijken onsets met een lage prominentie (lage sonorantie) klemtoon aan te trekken. We stellen voor om het gezichtspunt hier 180 graden te draaien en de hoge prominentie van de klinker *ten opzichte van deze onsets* de bepalende factor te laten zijn. Dus, alleen lettergrepen waarin het prominentie-verschil tussen onset en nucleus groot genoeg is overstijgen de gewichtsdrempel (zie boven). Met betrekking tot deze talen merken we verder op dat ze een zwak argument vormen voor het (indirecte) gebruik van onsets in klemtoonregels aangezien de data nogal omstreden zijn.

We besluiten hoofdstuk 5 met drie talen waarin eigenschappen van de onset die niet kunnen bijdragen aan de prominentie van de lettergreep toch de klemtoonplaatsing lijken te beïnvloeden. Weer zijn het Australische talen, namelijk het Mathimathi, het Ngarigu, en het Djapu. Een van de regels in het Mathimathi, bijvoorbeeld, plaatst klemtoon op de tweede lettergreep als die met een coronale consonant of met een klinker begint; anders valt de klemtoon op de eerste lettergreep. Zulke klemtoonregels zijn vreemd in elke metrische theorie, en we vermoeden dat er andere regels in het Mathimathi (en de twee andere talen) aan het werk zijn die de schijnbaar onset-gevoelige plaatsing van klemtoon veroorzaken.

In hoofdstuk 6 bespreken we het Mathimathi uitvoerig. Na een fonetische studie waarin het klemtoonpatroon van het Mathimathi (zoals dat in de literatuur wordt beschreven) wordt bevestigd, laten we zien dat dit patroon het gevolg is van een regel die het Mathimathi in een vroeger stadium heeft geleend uit de talen die in de buurt werden gesproken. Die regel schuift hoofdklemtoon op naar de tweede lettergreep. In het Mathimathi konden niet alle woorden aan deze regel gehoor geven; hoofdklemtoom wordt in het Mathimathi namelijk toegekend aan stammen, en sommige stammen zijn maar één lettergreep groot. Die stammen houden hoofdklemtoon op de eerste lettergreep, ook wanneer er suffixen worden toegevoegd (wat in bijna alle woorden het geval is). In stammen van twee of meer lettergrepen is er geen beletsel om de klemtoon naar de tweede lettergreep te schuiven. Wat nu het schijnbaar onset-gevoelige patroon veroorzaakt, is dat de mediale consonant van een tweelettergrepige stam in een ander proces verzwakt tot een coronaal of helemaal verdwijnt, zodat het achteraf lijkt alsof die coronaal de klemtoon heeft aangetrokken. In het laatste deel van hoofdstuk 6 werken we deze analyse volledig uit binnen het kader van de Optimaliteitstheorie.

In hoofdstuk 7 wordt het een en ander samengevat en besproken. We concluderen dat de aanname dat onsets gewichtloos zijn niet langer louter gebaseerd hoeft te worden op de afwezigheid van talen die onset-gewicht gebruiken. Er is nu ook fonetische evidentie voor de aanname dat onsets geen gewicht (kunnen) hebben. Wat de afwezigheid van onset-gewicht talen betreft kunnen we concluderen dat de meest klemmende tegenvoorbeelden, gehouden tegen het licht van hun historie, zich uitstekend lenen voor eenvoudige heranalyses die geen gebruik maken van onset-gewicht, zodat we dit type gewicht ook echt niet nodig hebben in de fonologie.

Curriculum Vitae

Rob Goedemans werd op 18 mei 1967 geboren te Haarlem. In 1983 behaalde hij het MAVO-diploma aan de Christelijke MAVO Uitermeer in Lisse, gevolgd door het HAVO- en het Atheneum-diploma, respectievelijk in 1985 en 1987, aan het Rijnlands Lyceum in Sassenheim. Na een jaar sterrenkunde te hebben gestudeerd in Leiden, stapte hij over naar Engelse taal- en letterkunde, waarvoor in 1989 de propedeuse werd behaald. Naast het voortzetten van de studie in de Engelse taal- en letterkunde startte hij met de kopstudie Algemene Taalwetenschap. In 1992 studeerde hij af in de Engelse taal- en letterkunde (met als specialisatie Fonologie), en in de Algemene Taalwetenschap in 1993 (met als specialisatie Fonetiek).

Op 1 september 1993 trad hij als Onderzoeker-in-Opleiding in dienst bij NWO. Het onderzoek dat heeft geleid tot dit proefschrift werd uitgevoerd aan het Fonetisch Laboratorium van de Rijksuniversiteit Leiden, onderdeel van de Faculteit der Letteren, en aan de Australian National University in Canberra, Australië.

Momenteel is hij werkzaam als Postdoc ten behoeve van een onderzoek naar de prosodie van Indonesische talen dat wordt uitgevoerd aan de Rijksuniversiteit Leiden, in samenwerking met de Universitas Indonesia in Jakarta. Bovendien werkt hij mee aan het toevoegen van geautomatiseerd oefen- en voorbeeldmateriaal aan een pas verschenen lesboek voor de Fonetiek.