

## Chapter 4

# Recursion in Phonology

### 4.1. Introduction<sup>18</sup>

In this chapter we investigate an instance of phonological recursion; more specifically, we investigate iterative rule application in phonological phrases. We will show that edge-marking processes, such as early pitch accent placement, can be applied recursively to phonological phrases that are embedded in larger phonological phrases. Before addressing prosodic recursion, we will demonstrate that recursion is quite a common phenomenon, not only in linguistics, but in nature, visual art, and music as well.

### 4.2. Recursion

Recursion can be seen as the repeated application of rules in generating a structure, in the sense that it applies to the output of every earlier application. In principle, there is no limit to the extension of the structure. It must be regarded as one of the most powerful mechanisms in generating complex systems, because it is possible to describe the whole system by describing just one layer of the system, since all layers are the same.

Recursion is claimed to be the only uniquely human component of the faculty of language (Hauser, Chomsky, and Fitch 2002). Animal communication systems allegedly lack this rich expressive and open-ended power of human language (Hauser, Chomsky, and Fitch 2002: 1570) that enables us to acquire a complex natural language on the basis of limited data. This statement is rather controversial. In fact Hauser et al. tone down their statement by suggesting that it is

---

<sup>18</sup> This chapter is an extension of Schreuder and Gilbers (2004a). It is supplemented with the acoustic results and an OT analysis.

possible that other animals may develop the same abilities, if recursion in humans evolved from the same cognitive capacity that is also used to solve other computational problems, such as navigation (p.1578). Indeed, examples of recursion can be found everywhere, as will be shown in this section.

Recursion is a very general principle. To begin with, all natural numbers are defined recursively (*cf.* Koster 2003), as shown in Table 24.

Table 24a Recursion in natural numbers

a	$1 = [1]$
b	$2 = [[1] + 1]$
c	$3 = [[[1] + 1] + 1]$
d	etc.

This same mechanism can be shown in the grammars of natural languages. For instance, each sentence can be embedded in a bigger sentence with similar structure, as shown in Table 24b:

Table 24b Recursive sentences

a	[he dreams]
b	[he dreams that [he dreams]]
c	[he dreams that [he dreams that [he dreams]]]
d	etc.

Each sentence can be expanded for ever by making it part of a bigger sentence with the same structure. There is no longest sentence. Mostly, such sentences will not be built of the same words entirely; the structures and types of words must be the same.

#### 4.2.1. Droste effect

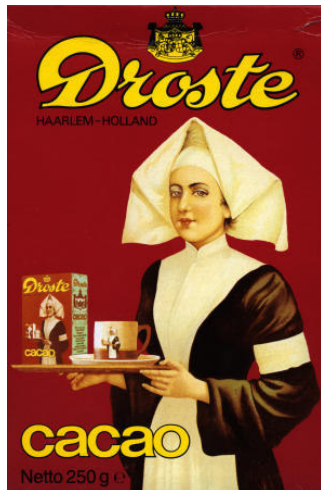
Everyone who uses a computer should be familiar with recursion. For example, the description of folders in a computer can be described by one rule: a folder can contain files as well as other

folders. This simple rule characterizes the recursive system of all folders in the computer.

The visual effect of recursion is shown in the two pictures in Figure 37. Figure 37a depicts the so-called ‘Droste effect’, named after a well-known Dutch cocoa tin. The Droste tin portrays a nurse with a tray on which we see a cup of cocoa and another Droste tin. This Droste tin portrays a nurse with a tray on which we see a cup of cocoa and another Droste tin, etc. Figure 37b depicts another example of visual recursion: the well-known Russian Matruska dolls. If you open the largest doll a smaller doll appears; if you open that doll a still smaller doll appears and so on.<sup>19</sup>

Figure 37 Visual examples of recursion

a Droste effect



b Matruska dolls




---

<sup>19</sup> In fact Figure 37b visualizes the Matruska dolls as non-recursive, because they are displayed next to each other.

#### 4.2.2. Fractals in nature

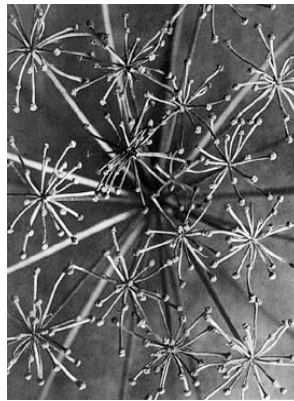
Structures that contain substructures with the same form as the whole, a property known as 'self-similarity', are also found in fractals (Mandelbrot 1977). This kind of recursion is apparent in nature and art. Just think of the leaf of a fern (Figure 38a): on each level the leaf consists of a central nerve with smaller leaves and each smaller leaf is built in exactly the same way. The pictures in Figure 38b by Karl Blossfeldt “Laserwort, part of a fruit umbel” and Figure 38c, a vegetable called romanesco, also show examples of recursion in nature. The umbel form and the spirally turreted shape are repeated in their extreme points.

Figure 38 Recursion in nature (see also Kawaguchi 1982, Smith 1984)

a leaf of a fern



b picture Karl Blossfeldt



c romanesco



#### 4.2.3. Endless loops in music and visual art

In 1981 the New Wave band The Look released an interesting single called “I am the Beat”. What makes this vinyl single interesting is the fact that it does not end. The final part of the groove is made circular and plays a rhythm that exactly matches the speed of 45 rounds per minute. That is why the final drum part of the song goes on and on, if the single is played on a non-automatic turntable. A similar circular groove was used by The Beatles on their B-side of the album “Sgt. Pepper’s Lonely Hearts Club Band” in 1967. The construction of this record, however, did not contribute to the construction of the song itself. Actually, this is an example of repetitive iteration.

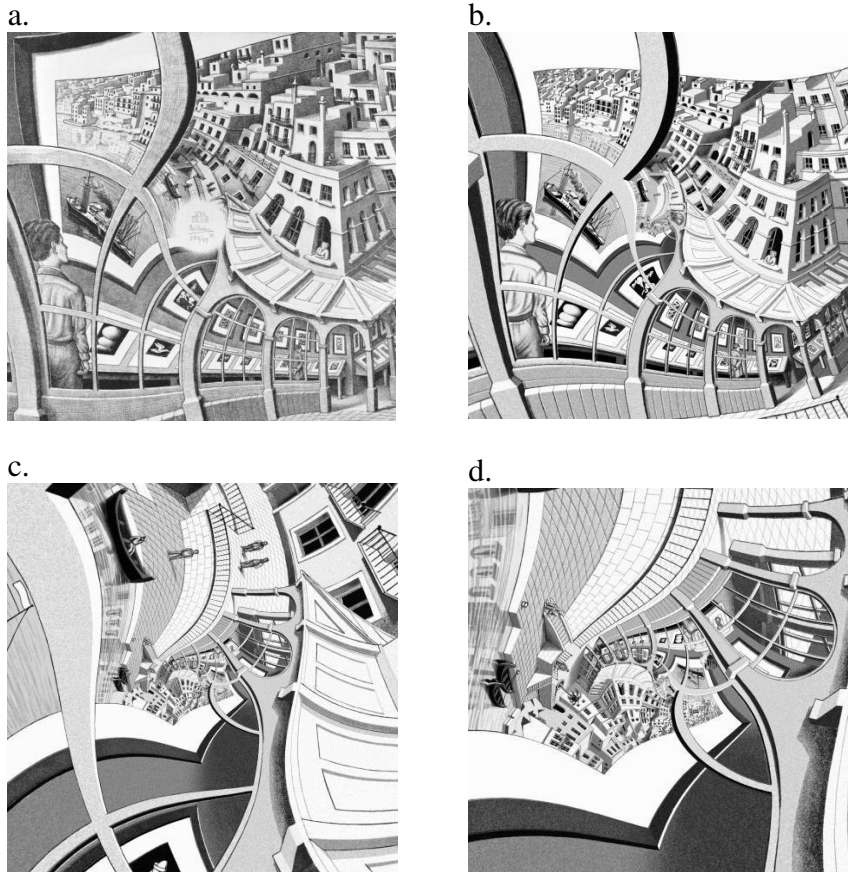
Hofstadter (1979) shows some examples of recursion in music in his book ‘Gödel, Escher, Bach’. An example of a piece that can be extended infinitely, is Bach’s ‘Canon per Tonos’ from ‘Das Musikalische Opfer’, which is also called the ‘endlessly rising canon’, because, from measure 6, it modulates in such a way that it climbs upwards one tone with every repetition. The canon as a whole is in the key C minor, but it concludes - or rather, *seems* to conclude - in D minor. And this “ending” connects smoothly to the beginning again, so one can repeat the whole canon in the key of D minor, which leads to E minor at the end, which again smoothly moves to the beginning. After six such modulations, the original key of C minor has been restored, but one octave higher, and the cycle can start all over again.

An endless loop that includes recursion is demonstrated in Escher’s “Prententoonstelling” (Print Gallery) (1956). In this well-known work a young man is viewing a print in a print gallery. On that print we see a Mediterranean seaport. The print is drawn with a certain elliptic curve, such that one of the buildings in the print, which happens to be a print gallery, at the same time forms the outside of the gallery in which the boy is standing. Escher himself was not able to draw the point where the gallery in the print merges into the big gallery and he therefore drew a white patch in the middle, which contains Escher’s monogram and signature (Figure 39a). In this central spot the recursion of the painting should have appeared.

The mathematicians De Smit and Lenstra (2003) from Leiden University solved the mathematical problem together with their

student Batenburg, and they reconstructed the structure to replace the white patch. Instead of Escher's white patch now a form of the so-called Droste effect appears: on the print in the print gallery we see another gallery with a young man looking at the print on which he sees the gallery with himself looking at... etc (Figure 39b). Thus, the spectator enters an endless loop. Figure 39c and d zoom in on the reconstruction of the white patch, where the recursivity is clearly shown. For animations of the Escher reconstruction, see De Smit and Lenstra's website [escherdroste.math.leidenuniv.nl](http://escherdroste.math.leidenuniv.nl).

Figure 39 Escher's 'Prententoonstelling' (1956) (a.), and its reconstruction (b., c., and d.)



#### 4.2.4. Recursively embedded structures in music

Another branch of art which is known to display recursive structures is music, although the term ‘recursion’ is not very commonly used. Recursion in music has many different manifestations. We find examples similar to fractals, which are referred to by the more general term ‘architectonic music’. In architectonic music large-scale structures echo small-scale aspects, particularly the organization of structures within structures. This is not an exceptional kind of structure in music. Some of these musical structures contain embedded phrases as well (Solomon 1998, see also his website for music examples: <http://music.theory.home.att.net/fracmus.htm>).

Some examples of recursive embedding can be found in Bach’s music. Hofstadter compares Bach’s ‘Kleines Harmonisches Labyrinth’ to a frame story. This piece is written as a labyrinth of quick key changes, in which the listener is soon disoriented. Normally tension builds up in a melody and finally resolves in the tonic, but here the modulations resolve in their own tonic, as if a story within a story ends, while the listener is still waiting for the ultimate resolution in the tonic. This means we hear the harmony recursively: we maintain a mental scheme of the keys, and each new modulation adds a new layer to the scheme. The listener keeps track of the overall key (the tonic) while listening to the local key (the pseudo-tonic) with its pseudo-resolution, and knows when the true tonic is regained. A perfectly recursive structure, center-embedding (*cf.* section 4.2.5), would be if the sequence of keys is retraced in reverse order.

The kinds of recursion in the examples above are not yet real examples of the embedded structures we are looking for. There is another canon of Bach, however, ‘Canon per Augmentationem in Contrario Motu’, canon 4 from the ‘Kunst der Fuge’, in which we find the embedded structure in a part of the melody. The note sequence in measures 30 and 31 contains a melodically identical copy of itself in measure 30, in a twice as fast tempo. The slower melody is in the bass voice, while the faster – embedded – copy is in the leading voice, as can be seen in Figure 40. The only difference is the second E in measure 30 in the leading voice, because a D would have clashed with the C sharp in the bass.

Figure 40 Recursion in Bach's *Canon per Augmentationem in Contrario Motu*

Whole phrases can be recursively embedded. Koch (1983) and Rothstein (1989) describe the phenomenon of phrase expansion, which is defined as the transformation of a phrase into a longer phrase, by adding more notes. These transformations are perceived as different representations of the same phrase. This phrase is the structural skeleton making up both phrases. In the experience of the listener a phrase expansion departs, often quite unexpectedly, from a fixed point of reference, and returns to it after a detour, bringing resolution and reassurance. The original phrase usually has a very regular hypermeter, while an expansion temporarily suspends this hypermeter without actually breaking it. Listeners can often “hear through” the expansion to the underlying hypermeter.

Not all phrase expansions are recursive in the sense we defined above. Rothstein (1989) distinguishes two kinds of phrase expansion that have recursive manifestations: external phrase expansions and internal phrase expansions.<sup>20</sup> External phrase expansions are an addition of subordinate material either before or after the basic phrase, leaving the basic phrase more or less unaffected. Internal expansions, however, add length within the basic phrase itself and are often literal or varied repetitions within the phrase. “Small” prefixes and “small” suffixes are external phrase expansions. A small prefix is less than a phrase; it is an incomplete phrase, an accompanying figure that sets the stage for a melodic entrance (Reicha 1814). This can also be found in vocal music, where often the melody is introduced by means of a short stretch of melody in the accompaniment just before a solo entrance, as a kind of “pre-imitation”. Measures 9 to 17 of the first movement of Schubert’s

<sup>20</sup> Rothstein does not mention the term ‘recursion’. The way he describes the phenomenon, however, is very similar to how recursion is defined above.

“Unfinished” Symphony, No. 8 in B minor, give an example of an expanded phrase with a small prefix (see Figure 41). The phrase begins with a prefix, which does not have a cadence of its own but instead moves to the accompaniment for the melody of the main phrase. The main phrase starts at measure 13. We can compare this kind of phrase to a recursive noun phrase in language, with two comparable adjectives: [*international [diplomatic organizations]<sub>NP</sub>*]<sub>NP</sub>; the phrase has two starting points, but only one end, which means that it constitutes a recursive phrase.

Figure 41 Recursive phrase: Schubert’s “Unfinished” Symphony, No. 8 in B minor, first movement, mm. 9-17

The musical score illustrates a recursive phrase structure in Schubert's "Unfinished" Symphony, No. 8 in B minor, first movement, measures 9-17. The score is in 3/4 time and B minor. It is divided into three systems. The first system (measures 9-11) is labeled "a) PREFIX" and "m. 9 Violins" with a *pp* dynamic marking. The second system (measures 12-14) is labeled "PHRASE:" and includes parts for Oboe (Ob.) and Clarinet (Cl.). The third system (measures 15-17) continues the phrase and ends with "etc." The structure shows a prefix that moves to the accompaniment for the melody of the main phrase.

Like prefixes, phrases can be expanded with suffixes. Suffixes are not forward-moving, developing structures; rather they form a stasis, an extension of a goal already reached (Riemann 1902). Small suffixes after a full cadence function as an expanded repose or codetta. A very common example is the expansion of a (full) cadence with another (full) cadence. The end of the suffix can serve as the



Figure 43 Haydn's Quartet in D minor, Op. 76, No. 2, 2<sup>nd</sup> movement, mm. 47-51; 61-62 (Rothstein 1989)

47

48

50

61

a)

b)

pizz

arco

V<sub>4</sub><sup>6</sup> - V<sub>3</sub><sup>7</sup> VI (deceptive cadence—long expansion leads to authentic cadence shown in b)

V<sub>4</sub><sup>(8)6</sup> - V<sub>3</sub><sup>7</sup> I

#### 4.2.5. Computing recursion

All examples mentioned are instances of regular context-free recursion, which is easily recognized by Finite State Automata (FSA). A different kind of recursion, center-embedding, is impossible to describe for FSAs, as the relationships between the embedding constituents on either side of the embedded constituent represent an unbounded quantity of information, which cannot be represented by finite states (*cf.* Chomsky 1959, Nederhof 2000). Center-embedding is found in human syntactic parsing, in sentences like *the rat [that the cat [that the dog chased] killed] ate the cheese* (Chomsky and Miller 1963). In theory people could pronounce such constructions, however, in practice people cannot store that much information in their working memory effectively till the verb of the sentence is parsed. We have not found evidence of such constructions in prosody thus far, whereas we could think of center-embedded musical structures (*cf.* section 4.2.4).

In the remainder of this chapter we will investigate whether the regular context-free kind of recursion can be found in prosody.

### 4.3. Recursion in phonology

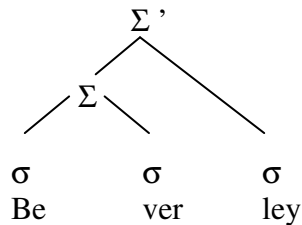
#### 4.3.1. Strict Layering and recursion

In the previous section we have shown many examples of recursion in nature, visual art, and music. Recursion has been called the ‘only uniquely human component of the faculty of language’ (Hauser et al. 2002). Hauser et al. make this statement in reference to syntax. Recursion refers to rules which are capable of repeated application in generating a sentence. In principle the number of prepositional phrases that may occur after a noun in a noun phrase is unlimited: *the American in the desert on a horse with no name*, in which *with no name* is a PP embedded in the PP *on a horse*, and you can always add a sentence to a sentence within a sentence as exemplified in the present sentence, and in Table 24b. As Crystal (1991, p.292) puts it, the application of recursive rules is the main formal means of

accounting for the creativity of language: by using this device, an infinite set of sentences can be generated from a finite set of rules.

In phonology, things seem to be different. Although iterative rule application is proposed for e.g. foot assignment, prosodic building rules seem to be limited in that sense. One cannot freely add e.g. onsets or nuclei to a syllable or syllables to a prosodic word. One of the rare occurrences of the incorporation of a prosodic domain within the same prosodic domain can be found in Selkirk (1980, 1984), who proposes a foot within a foot, constituting a super-foot, in order to account for dactylic patterns in rhythmic structures, as depicted in Figure 44.

Figure 44 Recursive foot (Selkirk 1980; 1984)

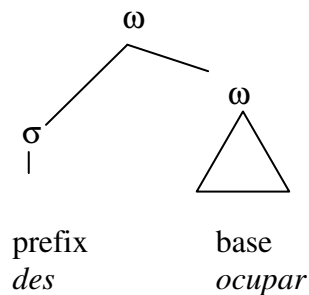


Selkirk's representation seems to give an argument for recursivity in phonological domains, although she circumvents recursion by calling the maximal foot a super-foot, indicating that both feet are of different categories. In her approach this kind of recursivity of phonological domains is restricted to the dactylic foot. Her approach is challenged by the fact that there are alternative representations without the need for recursivity. Dresher and Lahiri (1991) propose a ternary branching tree, and Kager (1994) maintains binarity and leaves the third syllable unparsed.

The limitations to the prosodic hierarchy are reflected in the Strict Layer Hypothesis (Selkirk 1984), of which one of the fundamental assumptions is that prosodic structure is not recursive. A mismatch thus exists between syntactically recursive constituent structure and the linearly segmented structure in prosody.

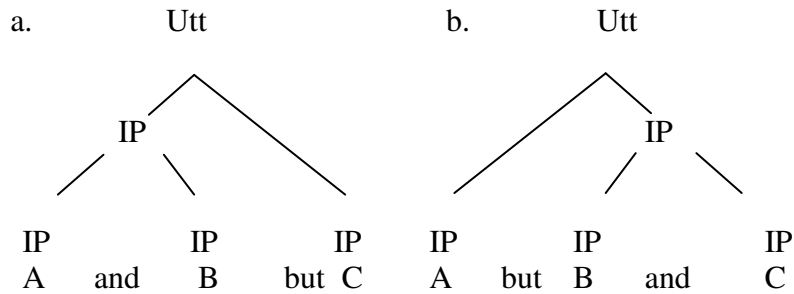
The Strict Layer Hypothesis has been criticized. Several phonologists, such as Itô and Mester (1992), have shown that in many cases it is violable; the assumption of non-recursivity has been challenged by a number of authors, in particular with regard to prosodic words with affixal clitics (Zec and Inkelas 1991, Booij 1996, Vigário 1999 among others). Vigário, for example, represents prefixed words in European Portuguese as a syllable and a prosodic word within a prosodic word, as depicted in Figure 45. Notice that similar prosodic domains should display similar phonological behavior. In European Portuguese word-internal vowels can be reduced, word-initial vowels cannot. Although the <o> in *desocupar* 'disoccupy' seems to be a word-internal vowel, it actually fails to undergo vowel reduction, because it is still the initial vowel of a prosodic word. This cannot be explained if recursion is not allowed in the prosodic hierarchy.

Figure 45 Recursive prosodic word (Vigário 1999)



Examples of recursion in larger prosodic domains like the Intonational Phrase can be found in Ladd (1986, 1996), and references therein. The Intonational Phrase (IP) is the domain of a perceptually coherent intonational contour (Shattuck-Hufnagel and Turk 1996). These Intonational Phrases are dominated by the Utterance (Utt), which is often described as the phonological counterpart of the syntactic sentence. It is described as the largest span of application of phonological rules (Selkirk 1978, 1980, Nespor and Vogel 1986, Hayes 1989).

Figure 46 Recursive intonational phrase (Ladd 1986, 1996)



Ladd represents different kinds of conjunction as in Figure 46. Phonetic differences, such as pause-duration cues and declination slope reset motivate the division into three Intonational Phrases, yet the *but*-boundaries are significantly stronger than the *and*-boundaries. The difference in strength between the conjunction of (*A and B*) as opposed to (*A and B*) *but* *C* is visualized in Figure 46 by embedding (*A and B*) in a dominating IP. Without recursion, we cannot explain the phonetic differences between the two kinds of conjunction. The following sentence, which has the structure of Figure 46a, exemplifies these different conjunctions: *Aretha Franklin is a soulful singer, and Carole King is an excellent composer, but Florence Foster Jenkins amuses us the most.*

Similar arguments to the ones above can be found in Ladd (1992), Inkelas (1989) and McCarthy and Prince (1993a,b). These arguments led Selkirk (1995b) to replace the Strict Layer Hypothesis with a series of four separate constraints, one of which is Non-Recursivity: No  $C_i$  dominates  $C_j$ ,  $j = i$ . This is a violable constraint, in terms of standard OT (Prince and Smolensky 1993).

Since recursion is very common in syntax, the primary source of evidence for instances of recursion in phonology is probably provided by phonological rules that operate over (morpho-) syntactically defined recursive domains. Notably, in the phonological domains where morphosyntax does not play a role, no

evidence has been found for recursive structure - like the syllable domain - or the arguments are weak, like for the domain of the foot. The phonological phrase is one instance of a domain of which the phrase breaks typically coincide with the edges of morphosyntactic phrases (Selkirk 1984, Nespor and Vogel 1986). Although there is no consensus on what exactly constitutes the phonological phrase, we follow Selkirk, who assumes that the phonological phrase aligns with either the left or the right edge of the head of a maximal projection which is not lexically governed, i.e. it groups a phrasal head together with its adjacent modifiers and functional elements (Selkirk 1995b).

#### 4.3.2. Research question

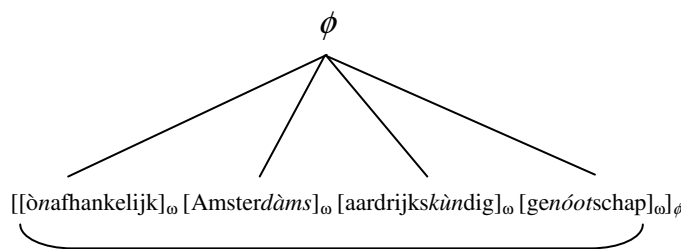
In this chapter we will show that recursion can be found at the higher domains in the prosodic hierarchy. To investigate this, we examined Dutch phonological phrases, consisting of either a noun modified by one adjective, of the type *aardrijkskundig genóotschap* 'geographical society', or by two adjectives, of the type *Amsterdàms aardrijkskundig genóotschap*, i.e. a recursive noun phrase (the accents on the adjectives indicate the main stress position in citation form.). Syntactically, this kind of phrase can in principle be infinitely extended with adjectives: [*onafhankelijk* [*Amsterdàms* [*aardrijkskundig genótschap*]<sub>NP</sub>]<sub>NP</sub>]<sub>NP</sub> 'independent Amsterdam geographical society'.

The first type of phonological phrase, with one adjective, is known to display early pitch accent placement (Shattuck-Hufnagel 2000, Shattuck-Hufnagel et al. 1994) as a means of signaling a phrasal boundary to give the listener a cue to the prosodic structure of the spoken utterance. This phenomenon was first referred to as Iambic Reversal (Lieberman and Prince 1977) and it is also known as stress shift, the (English) Rhythm Rule (Lieberman and Prince 1977), or the Phrasal Rule (Hayes 1984, Shattuck-Hufnagel 2000). In contradistinction to these stress movement accounts, Gussenhoven (1991) proposed that the phenomenon is not movement of lexical main stress, but a combination of two events: the occurrence of a phrase-level intonational prominence on the earlier full-vowel

syllable, and the non-occurrence of a pitch accent on the later main-stress syllable. Horne (1990), Grabe and Warren (1995), Shattuck-Hufnagel et al. (1994), and Vogel, Bunnell, and Hoskins (1995) subsequently showed that that indeed was the case. We can therefore say that Hayes' Phrasal Rule is a boundary marking phenomenon, which marks phrasal boundaries. As we will show, however, the boundary marking not only depends on pitch, but also on duration and timing (*cf.* section 4.4.5.2). Moreover, Quené and Port (2003) found that rhythmic timing also has a strong influence on the process.

The question now is what kind of prosodic structure has to be assumed for the second type of phrase, modified by two adjectives, the syntactically recursive noun phrase. If the non-recursivity assumption holds and these phrases are non-recursive, then they must have a flat, linear structure, and no early accent will occur on the intermediate adjectives: [*ònafhankelijk Amsterdàms aardrijkskùndig genóotschap*]<sub>φ</sub>, as depicted in Figure 47.

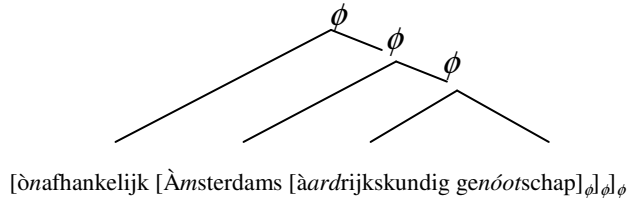
Figure 47 Non-recursive phonological phrase



If, on the other hand, an early accent can be realized on these intermediate adjectives, we have to assume a nested, recursive phrase structure, indicating two or more left boundaries within the same phonological phrase, i.e. a nested prosodic structure: [*ònafhankelijk*

[*Àmsterdams* [*àardrijkskundig genóotschap*]]], as shown in Figure 48.<sup>21, 22</sup>

Figure 48 Recursive phonological phrase



To investigate whether these kinds of phonological phrases can indeed be produced with a recursive prosodic structure, with the Phrasal Rule applying two or more times, we conducted an experiment, which is described in the next section.

#### 4.4. The experiment

##### 4.4.1. Task design

In order to get as close to spontaneous speech as possible, we used the Map Task (Brown et al. 1984) to build our corpus in a controlled way. The Map Task is originally a cooperative task involving two participants, used to build dialogue corpora. We adapted the original design somewhat to our own requirements. The subject and the experimenter sat opposite one another, the subject sat in the soundproof studio behind a glass window, and each had a map which the other could not see. The subject had a map consisting of a starting

<sup>21</sup> Selkirk (1995a) herself gives some examples of a similar kind in English: [[nòrthern][Càlifornia wínes]] as opposed to the right-branching phrase [[nòrthern Califòrnia] wínes], but without going into its recursivity. We hypothesize that these syntactically recursive noun phrases can be realized as recursive phrases in prosody as well.

<sup>22</sup> In terms of computational linguistics, this is a regular context-free right-recursion, which is quite easy to describe for finite-state automata (*cf.* Chomsky 1959, Nederhof 2000).

point, an endpoint and some landmarks, labeled with their names, on the route. The phrases of interest were two of the landmarks, the rest were fillers (see Figure 49). The experimenter's map only had the starting point drawn on it, which makes the experiment more or less 'double blind'.

We made fifty different maps, with two landmarks of interest on each map, which makes a hundred phrases in total. Each map had four fillers and the phrases of interest never appeared as a starting point or endpoint. Each map contained one syntactically recursive phrase landmark [Adj [Adj Noun]] and one non-recursive, non-corresponding phrase landmark [Adj Noun]. The subjects never saw two corresponding phrases.

#### 4.4.2. Subjects

We tested 24 subjects, ten men and fourteen women, aged 19 to 28. Most of them were law students, with Dutch as their mother tongue. Ten subjects were brought up in the northern provinces of the Netherlands, nine of them came from the center, three from the west and two from the south. One subject had grown up in the Netherlands Antilles, and Dutch was not her mother tongue, though she learned it in her childhood. We found no differences in the characteristics of interest, so we kept her in the experiment. We did not find any regional influences on the results either. The subjects were unaware that it was a linguistic experiment.

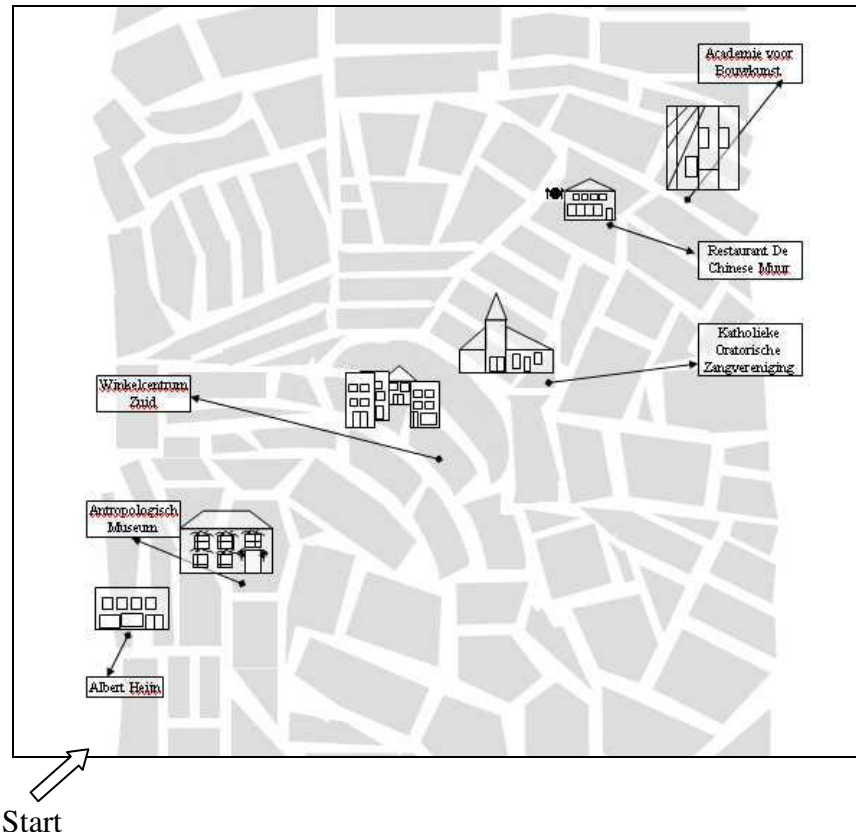
#### 4.4.3. Method

The subjects were told that their goal was to lead the experimenter from point A to point B on the map, leading past all landmarks on the map, and they were supposed to mention all the landmarks they came across. The picture in Figure 49 shows part of such a map. The following text is a translated fragment of a map description of the map in Figure 49 by the male Subject M04. (Bold phrases are the English glosses of the landmarks depicted on the map.)

Question: *'We are at **Albert Heijn's** and I want to go to the **Academy of Architecture**. Can you tell me the way?'*

Answer: *'Ehmm, - let's see - in front of **Albert Heijn's** – you turn left, – and then you take eeh – then go straight ahead and then you turn eeh – second street left, – annd, then just walk straight ahead, then you come across the **Anthropological Museum**, [...] – eehm – when you stand in front of the **Anthropological Museum** you just walk straight on, – annd eehmm – then you come across eeh first a T-junction then you keep on walking straight ahead, at the crossroad you just cross it, – eehm – then you again arrive at a T-junction – kind of a T-junction – then turn right, – annd eehmmm – [...] then you follow a slight curve to the right, eehmm, [...] after that you turn to the first on your left immediately, then you arrive at the **Shopping Centre South**, eehhm – then immediately to the right again, annd – tooooo the left, [...] ehm you walk straight ahead, then you come across the church of the **Catholic Oratory Choral Society**, eehm arrived there ehm you turn right, annd after that immediately to the left and then you walk up that eeh that street that you were walking already – [...], and then you arrive at the **Academy for Architecture** and before that you came across **Restaurant the Chinese Wall**.'*

Figure 49 A map used for the map-task



The experimenter did not interfere. Afterwards the subjects were asked to read the adjectives aloud in citation form, within the sentence *Ik spreek nu het woord ... uit* 'I now pronounce the word ...'.

All data were recorded with a Sennheiser MKH 40 Microphone (mono), on a Sony DTC-57ES DAT-recorder, with Fuji Digital Audio Tapes. The sound files were digitalized with Cool Edit Pro at a 22050 Hz sample rate, mono with 16-bit Resolution, normalized to 100%, and saved as .wav files (Windows PCM). The phrases of interest were extracted from the sound materials; the same procedure was used for the citation form words.

For the auditory analysis five trained listeners judged the data auditorily and indicated on which syllables in the adjectives they perceived word accent. They were free to indicate more than one accent per adjective, which meant that words could be double pitch accented. A majority judgment of the five trained listeners was decisive; it turned out that there was consensus among three listeners on most data. We ran Chi-Square tests on the statistics.

For the acoustic analysis we analyzed the data in PRAAT (Boersma and Weenink 1992-2006). We measured fundamental frequency in Hz (maximum and mean), duration in seconds and intensity in dB of the rhymes of the two syllables of interest, for each phrasal adjective. Furthermore, we measured spectral balance in a small set of the data and we measured rhythmic timing between the perceived accents in the phrases. The same measurements were made on the words in citation form, and we compared the values of the words in phrases and in citation form with T-tests. A Multivariate Analysis of Variance showed us which acoustic cues were responsible for the perception of a pitch accent, and moreover, we did Chi-square tests to compare the values of the words in phrases and in citation form in relation to the perceived accents. Inter-accent intervals were compared with T-tests as well.

#### 4.4.4. Data

As pointed out above, the data consisted of one hundred phonological phrases, half of which were [Adjective Noun] combinations and the other half corresponding [Adjective [Adjective Noun]] combinations. Table 25 shows a selection of our data.<sup>23</sup>

---

<sup>23</sup> Some examples from the experiment can be downloaded as MP3-files from <http://home.planet.nl/~schre537/sounds.htm> or [www.maartjeschreuder.nl](http://www.maartjeschreuder.nl).

Table 25 Data

<i>Aardrijkskundig genootschap</i>	'geographical society'
<i>Amsterdams aardrijkskundig genootschap</i>	'Amsterdam -'
<i>Diplomatieke organisaties</i>	'diplomatic organizations'
<i>Internationale diplomatieke organisaties</i>	'international -'
<i>Regionale dagbladders</i>	'regional daily press'
<i>Algemene regionale dagbladders</i>	'general -'
<i>Socialistische partij</i>	'socialist party'
<i>Progressieve socialistische partij</i>	'progressive -'
<i>Psychiatrisch ziekenhuis</i>	'psychiatric hospital'
<i>Academisch psychiatrisch ziekenhuis</i>	'university -'

In order to minimize the influence of pure regular rhythm instead of prosodic structure, for example eurhythmicity effects of the Quadrisyllabic Rule (Hayes 1984), we varied the number of syllables between the accentable positions in the words from 1 to 7.<sup>24</sup> We also avoided stress clash effects, and contrast effects of phrases ending in a similar suffix or contrasting phrases on one map.

The phrases we used are of a type which can undergo so-called Rhetorical Retraction (Gussenhoven 1983, 1984). Rhetorical Retraction in fact refers to the same phenomenon as early pitch accent placement, but Van Bezooijen (2001) shows that speakers use Rhetorical Retraction mostly as a propagandistic speech style. Gussenhoven (1983, 1984) shows that the effect of style is significant but small, in the least rhetorical style. In our experiment we made the speech context as neutral as possible, in order to show that early accent placement is not only a stylistic phenomenon, but also a structural device. With the map task we were guaranteed a non-commercial, neutral context, and we will not use the term Rhetorical Retraction for our data.

<sup>24</sup> Recall that we counted the Phrasal Rule, which is one of Hayes' original eurhythmicity rules, as a boundary marking principle, no longer as a (eu)rhythmicity rule.

Ten out of the fifty recursive NPs can be interpreted as if the first adjective modifies the second, instead of the noun, as possible in e.g. *progressief individualistisch verbond* ‘progressive individualistic union’. If subjects indeed interpreted a phrase as ‘progressively individualistic union’, the recursive prosody was not expected to be realized in that phrase. We will look into the data whether phrases with such an ambiguous interpretation behave differently from the rest.

We divided the subjects over five different map sets, so each subject read ten maps, which means ten recursive and ten non-recursive phrases for each subject. This resulted in about 550 spoken phrases in total. An impressionistic observation reveals that the subjects mostly pronounced the names of the landmarks in focus, and most of the times it was before a comma with a pause, pronounced with a so-called continuation rise (L-H%) or before a full stop, with the so-called declarative contour (L-L%) (Pierrehumbert 1980, Gussenhoven et al. 1999). Some subjects occasionally repeated the phrases. The repetition was then most of the time out of focus and sometimes with a different rhythmic pattern. Others unfortunately missed some of the phrases. In the results section we only report on single, first uttered, utterances of a phrase for each subject.

#### 4.4.5. Results

##### 4.4.5.1. *Auditory results*

Five trained listeners indicated where they perceived pitch accents on the adjective in the non-recursive noun phrases and on each of the two adjectives in the recursive noun phrases. When a majority indicated they perceived a pitch accent on a certain syllable, this syllable was appointed a 1, the other potential pitch accent site was assigned a 0. In most of the cases the listeners agreed on the pitch accent position as either 0 or 1. Figure 50a shows the percentages of perceived pitch accents perceived on the early syllable, on the main stress position, and also the percentage of the cases that were judged to be double accented. Figure 50b depicts the results for the words in

citation form, where the main stress position is accented almost always.

The graph in Figure 50a clearly shows that, although there is a strong preference for the subjects only to accent the main stress syllable, early accent placement is also a strong tendency. For the non-recursive phrases almost 30% displayed an early accent. This is the same percentage Van Bezooijen (2001) reports for rhetorical expressions, while we had a totally neutral context. This already is a surprising result.

The most interesting result for this study, however, is the early accent bar of adjective 2 of the recursive phrases in Figure 50a. Although these adjectives were not the initial words of the longer phonological phrases, they still received an early accent in 22% of the phrases. This result seems to confirm our hypothesis that these syntactically recursive phrases can also be recursive prosodically (Schreuder and Gilbers 2004a).

Subjects and items differ greatly in their behavior and patterns, however. The standard deviations are very large or even maximal for the items. This means that some items never conformed to the Phrasal Rule, and others, on the other hand, always did. We could not find any systematic characteristic in the items which did or did not show many shifts. It seems not to be a rhythmic phenomenon: the number of syllables between accentable syllables in a phrase does not influence the number of early pitch accents on that phrase. Gussenhoven (1983), on the other hand, found that if one of the interstress intervals contains a low number of syllables, the propensity for early accent placement increases.

Another interesting finding is the relatively high percentages of double pitch accents on phrasal adjectives, i.e. a pitch accent both on the early syllable and the main stress position of the word. Shattuck-Hufnagel (2000) reports a similar finding. These two findings are an argument for the view that it is not a matter of stress shift, but a phrasing phenomenon, which involves accents. This means that the eurhythmy rules (Hayes 1984) do not apply to these data, at least not when we base rhythm on syllable counting.

As said in section 4.4.4, some of the recursive NPs have a possible interpretation in which Adjective 1 modifies Adjective 2. In

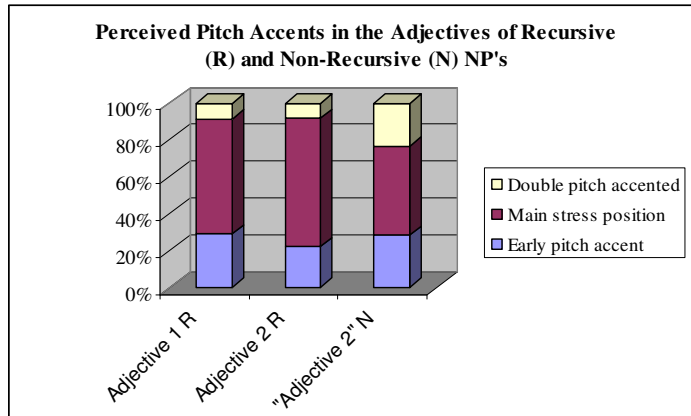
the data we see quite a large amount of non-restructured, and non-recursive realizations with an early accent only on Adjective 1, but also a number of recursively early pitch accented ones. This means the observations for these phrases do not differ much from the overall pattern.

Besides the subject and item dependencies, Pearson Chi-Square tests show that the proportions of early accents and main stress positions in the phrase and in citation form are significantly different ( $\chi^2$  (df 2) = 117.209,  $p < 0.001$  for the non-recursive NPs,  $\chi^2$  (df 2) = 54.552,  $p < 0.001$  for the recursive NPs). The adjectives in citation form had a nearly 100% score of pitch accents in main stress position. This means there is a substantial proportion of the data in which an early accent is perceived on adjective 2, both for the non-recursive and the recursive phrases. However, the difference between the proportions of corresponding adjectives of the non-recursive and recursive phrases is also highly significant ( $\chi^2$  (df 2) = 20.393,  $p < 0.001$ ). Note that the Chi-Square value of the non-recursive vs. recursive test is much smaller than the values of the phrase vs. citation form tests, which indicates that the difference between the patterns of the corresponding adjectives in phrase-initial and phrase-second position are much smaller than the differences between the adjectives spoken in phrase-second position and in citation form. Nonetheless, we can conclude that listeners do perceive prosodic recursion.

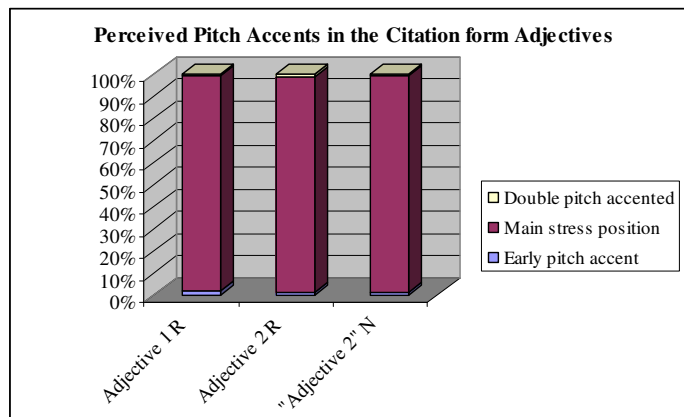
In the next section we try to find acoustic evidence for the perceived early pitch accents.

Figure 50 Percentages of perceived pitch accents on the adjectives in the Recursive phrases (R) (N = 232) and Non-Recursive phrases (N) (N = 238)

a.



b.



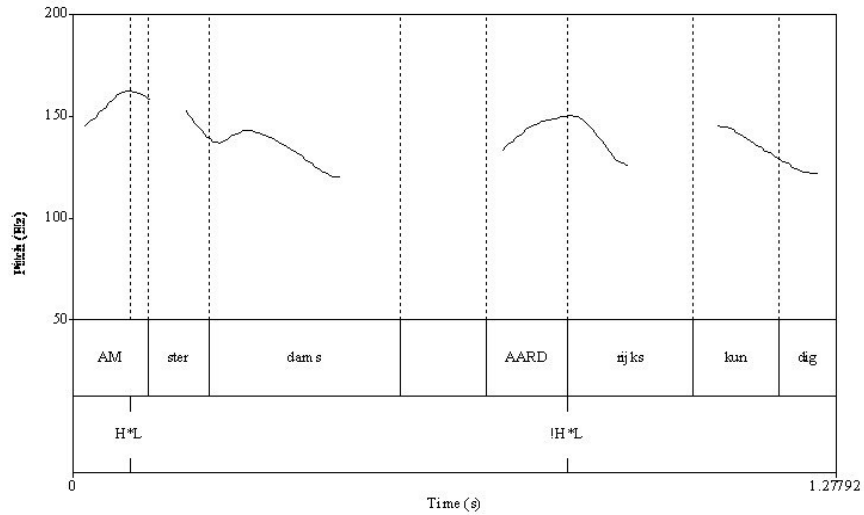
#### 4.4.5.2. Acoustic results

In order to try to underpin our findings with empirical evidence, we measured the values of the correlates maximum and mean fundamental frequency, duration, intensity, spectral balance, and also rhythmic timing.<sup>25</sup> For accent, measuring fundamental frequency actually should suffice, but we did not want to exclude other possible correlates in advance. Figure 51 shows the F0-contours and syllable durations of *amsterdams aardrijkskundig (genootschap)* in the phrase (Figure 51a) and in citation forms (Figure 51b), as realized by the male subject M05. This phrase was perceived as having early accents on both adjectives, whereas the words in citation forms were accented on the main stress position. This is reflected in these F0-contours. The accented syllables are capitalized in the textgrids.

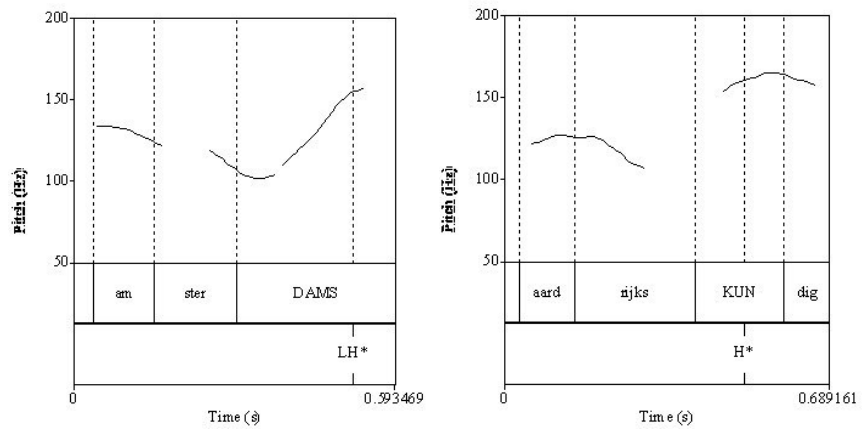
---

<sup>25</sup> For the recursive NPs the analyses of pitch, duration, and intensity were performed only on the second adjectives, which were the same adjectives as in the non-recursive NPs. Rhythmic timing was analyzed separately from the more conventional correlates. We will report on these rhythmic timing analyses further on in this section.

Figure 51a. F0-contour of the phrase *àmsterdams àardrijkskundig (genóotschap)*



b. F0-contours of the words *amsterdáms* and *aardrijkskúndig* in citation form



In order to look for the correlate of perceived accent in the phrases, we performed a Multivariate Analysis of Variance (MANOVA) on the phrasal adjectives. We investigated whether the

perception of an accent on a certain syllable depended on the height of (one of) the acoustic correlates duration, fundamental frequency, intensity or spectral balance of that same syllable, either as a bundle of correlates, or as a single correlate. This turned out mostly to be true for the main stress syllable, however not for the early syllable. T-tests comparing the phrasal correlates with those of the individual words also showed no significant differences for the early syllables, but some significant differences for the main stress positions.

These first findings suggested that we should shift our focus from the acoustics of the early syllable to the influence of the main stress syllable acoustics on the perception of the early syllable. Therefore we performed another MANOVA, but now with the perceived accent on the early syllable compared with the acoustic values of the main stress syllable. Indeed, where an accent was perceived on the early syllable, the acoustic correlates of the main stress syllable show lower mean values, except for intensity, as can be seen in Table 26 and Figure 52. The correlates that do show lower main values are fundamental frequency (pitch) and duration. Notice that the high standard deviations for the fundamental frequency are partly caused by the fact that we analyzed male and female speakers together.

This may seem a strange finding, meaning that listeners perceive a change on one syllable, while the change in reality had occurred on another syllable. However, Horne (1990), Grabe and Warren (1995), and Vogel, Bunnell and Hoskins (1995) also find evidence for accent deletion on the basis of phonetic data. The main stress position under stress shift represents a deleted pitch accent, while the secondary stress is unchanged. The finding in this chapter of course strongly supports this view. Gussenhoven (1991, 2005) accounts for non-peripheral accent deletion in phonological phrases phonologically. According to his proposal the distribution of pitch accents in English is determined by the interaction of constraints on lexical accent working on two lexical levels, and postlexical rhythmic readjustments (*cf.* section 4.4.5.2).

The fact that the perception of an accent on the early syllable depends on the deletion of an accent on the main stress syllable, is indicative for the relativity of our perception: the removal of an

accent on one syllable has the effect of a perceptive accent on another syllable, while no acoustic accent cues may be present on that syllable. Table 27 shows the result of the MANOVA: all the differences but intensity are significant.

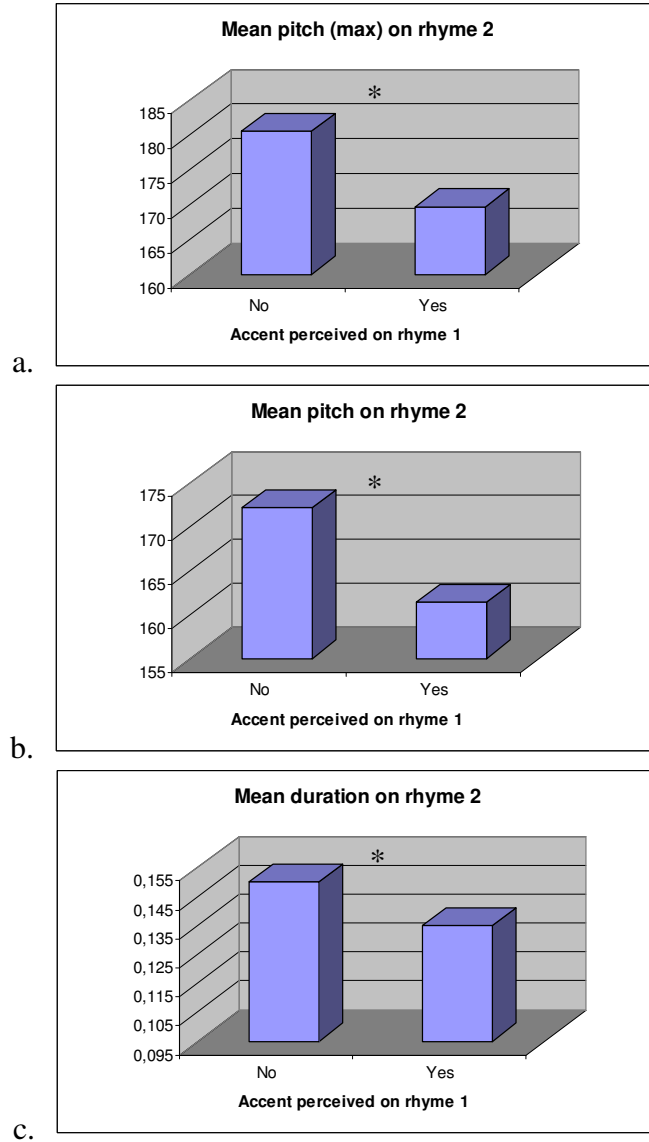
Table 26 MANOVA: Descriptive Statistics

Acoustics on main accent position	Accent perceived on early syllable	Mean	Std. Dev.	N
pitch max	No	180.5941	52.87693	266
	Yes	169.6874	46.29735	184
pitch mean	No	172.1913	49.80910	266
	Yes	161.4167	44.54887	184
duration	No	.1500	.06776	266
	Yes	.1357	.05071	184
intensity	No	72.2209	5.07711	266
	Yes	72.1551	6.08775	184

Table 27 MANOVA: dependencies between perceived early accents and absence of acoustic accents on the main accent position

Source	Dependent Variable (acoustics of rhyme 2)	F	Sig.
Perceived accent on rhyme 1	pitch max	5.115	.024
	pitch mean	5.542	.019
	duration	5.869	.016
	intensity	.016	.901

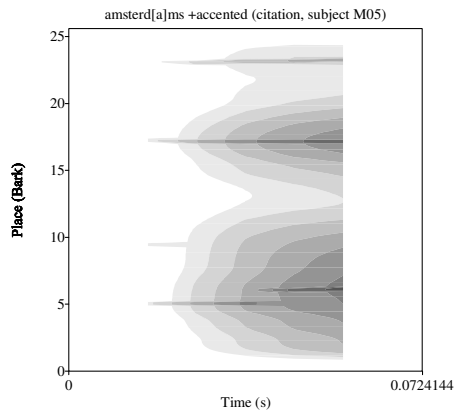
Figure 52 Acoustics on main accent position



We also performed an impressionistic investigation on spectral balance in a small part of the data. (For the description of this correlate, see Chapter 3.) It appeared that spectral balance did not contribute to the effect of accent. In Figure 53a,b,c we show the cochleagrams and the loudness distribution over the spectrum of the main stress syllable of the word ‘amsterdams’ with an accent on that syllable in citation form, and without an accent in phrasal form. The spectral balance in these pictures shows the expected higher energy in the higher frequency regions in the accented case. However, in most other cases the differences were unpredictable. So we assume spectral balance not to contribute to the perception of (early) accent.

Figure 53 Spectral balance on main accent position

a. Cochleagram of +accent [ɑ]



b. Cochleagram of -accent [ɑ]

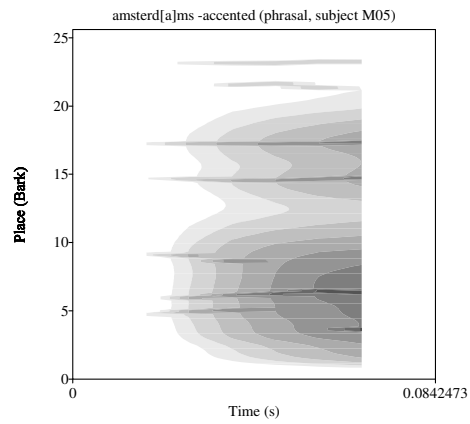
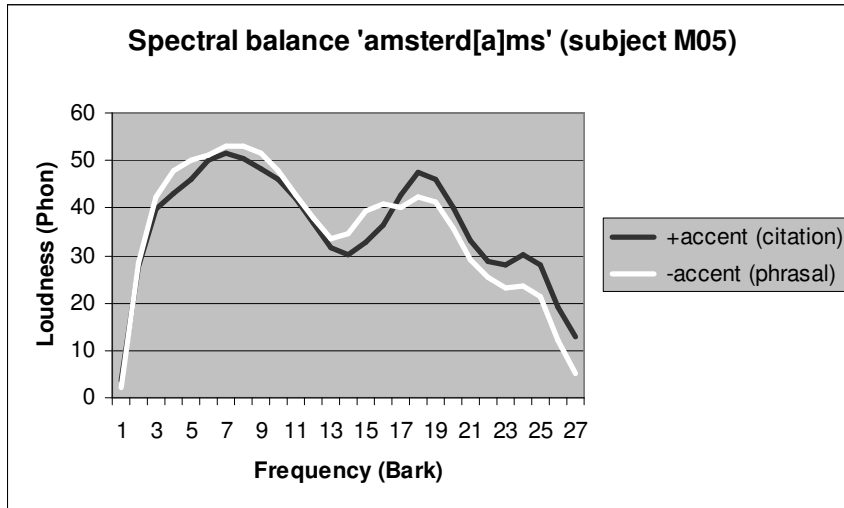


Figure 53c. Loudness within the spectrum



Subsequently, we were interested in the differences between the main stress syllables of the phrases and the words in citation form. We subtracted the values of the correlates of the citation form adjectives from the values of the corresponding phrasal adjectives. An outcome smaller than zero meant that the main stress syllable of the citation form word had higher values than those of the phrasal words, hence the main accent syllable of the phrasal word should have been de-accented. Outcomes smaller than the Just Noticeable Differences (Rietveld and Van Heuven 1997 and references therein) are counted as having equal values in citation form and in the phrase.

With a Pearson Chi-Square test we compared the outcomes of the main stress syllables with the perceived accents on the early syllables. The results from our pilot experiment (138 items) showed that the de-accenting of the main stress syllable did in fact result in the perception of an accent on the early syllable. This is particularly true for the correlates mean pitch ( $\chi^2$  (df2) = 13.056,  $p < 0.001$ ) and duration ( $\chi^2$  (df2) = 6.891,  $p < 0.05$ ).<sup>26</sup> Maximum pitch gives just a

<sup>26</sup> Note that many words that were not perceived as early accented also have low values on the main stress syllable. This seems to contradict the other findings. Possibly, this is due to the fact that the early syllables of many of the words in

marginally significant result and intensity rather does the opposite from the other correlates and the dependency is not significant. Intensity may be not a reliable correlate, due to testing circumstances like the variable distance of the subject's mouth to the microphone. The dependencies between accent and the correlates mean pitch and duration are plotted in Figure 54. The data of the full experiment did not confirm these last findings with significant results. Nonetheless, the mean values of the phrasal data alone do point to the dependency between the acoustics of the main stress syllable and the perception of an accent on the early syllable, as shown above.

As mentioned in the introduction of this section, we also analyzed rhythmic timing. Quené and Port (2003) found that the perception of early pitch accents depends on equal spacing, i.e. rhythmic timing. Note that this is a different kind of rhythm than counting syllables as in the eurhythmy rules. We looked for the same evidence in a subset of our data. First we compared, by means of a t-test, the two inter-accent intervals between the perceived early accents on adjectives 1 and 2 and the main stress accent on the noun. These were found to be similar ( $t(df30) = -.307, p > 0.05$ ). Figure 54a shows the equality of these two intervals, as compared to the intervals between the non-accented syllables in adjectives 1 and the early accented ones in adjectives 2.

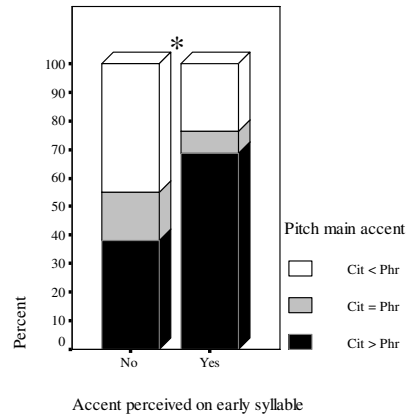
This points to the same conclusion as Quené and Port's (2003), that the perception of early pitch accents is caused by rhythmic timing. However, in the foregoing paragraphs we have shown that, as opposed to our data in the fast speech experiment (chapter 3), the perception of an accent does not entirely depend on rhythmic timing alone, because the acoustic correlates pitch and duration of the main stress position – de-accenting - also influence the perception of early accents.

---

citation form have relatively high acoustic values, just not high enough to be perceived as an accent, relative to the main accent syllable. These high early syllables reflect the fact that words in citation form are phonological phrases by themselves. Normally, phrases consisting of a single word are not signaled by early accent placement, but nevertheless this may have caused the high 'Cit > Phr'-bar in the chart in Figure 54 where no early accent was perceived in the phrases.

Figure 54 Acoustic values Pitch and Duration of the words in citation form subtracted from the values of words in a phrase (pilot experiment, 138 items)

a.



b.

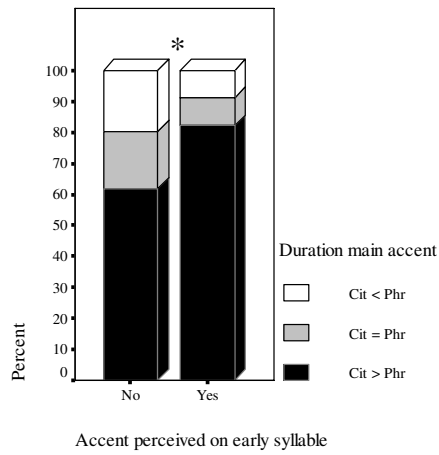
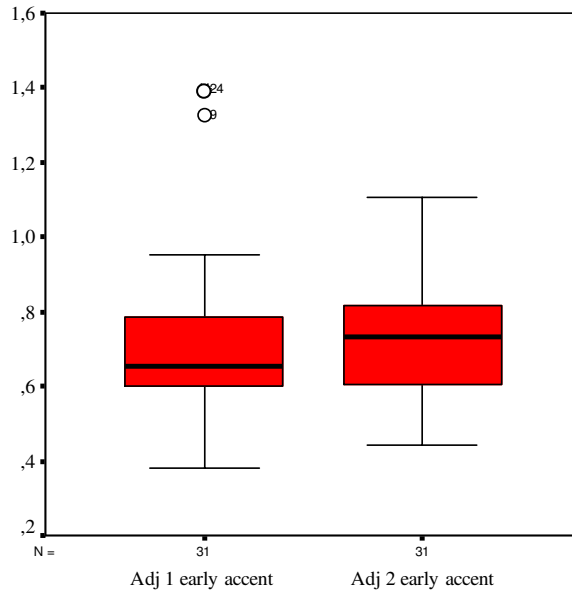
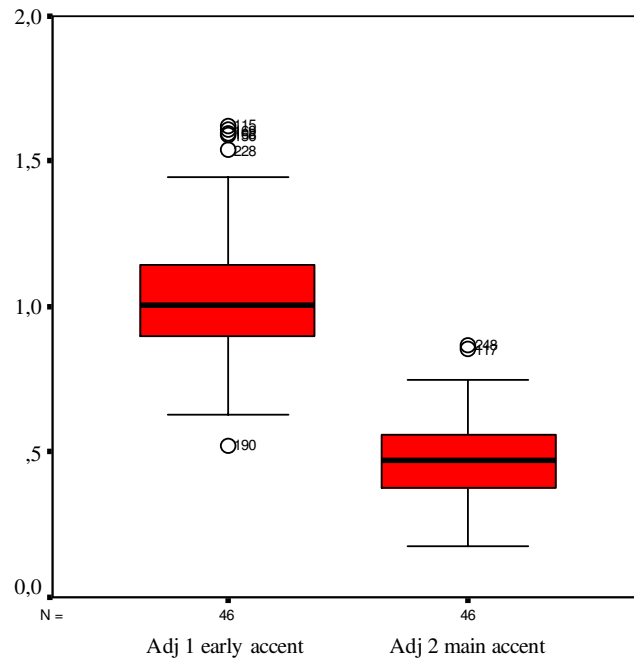


Figure 55a Boxplot of inter-accent intervals between the early accents on adjectives 1 and 2.



Moreover, if timing caused the perception of accents, this should also hold for non-shifted accents. We compared the inter-accent intervals between the early accented syllable of adjective 1 and the accented main stress position of adjective 2 on the one hand, with the interval between this last accent and the accent on the noun on the other. These intervals appeared to be significantly different ( $t(df44) = 13.487, p > 0.001$ ), as the boxplot in Figure 55b shows. We must conclude that more is involved than rhythmic timing only. On first sight, the Equal Spacing Constraint seems to have a strong influence on the perception of early accents, but we showed that this is not the case between early and non-shifted accents.

Figure 55b Inter-accent intervals between the early accents on adjective 1 and the accents on main stress position on adjective 2



#### 4.4.6. Phonological analysis

For the phonological analysis of the data, we will give an OT analysis, with a fixed constraint ranking. The optimal output must be the most observed candidate. Suboptimal alternatives, however, can emerge as variants. As in the previous chapter, variation forms are modeled as local optima with respect to a neighborhood structure on the set of candidates, in terms of the optimisation algorithm of Simulated Annealing (Bíró 2005, to appear). In the previous chapter we described Bíró's Annealing Simulation for our fast speech data, and this simulation proved to be successful. We did not perform an Annealing Simulation for the current data, but it is possible that a

simulation would come up with similar percentages (Bíró, personal communication). It should be tested in future work, however.

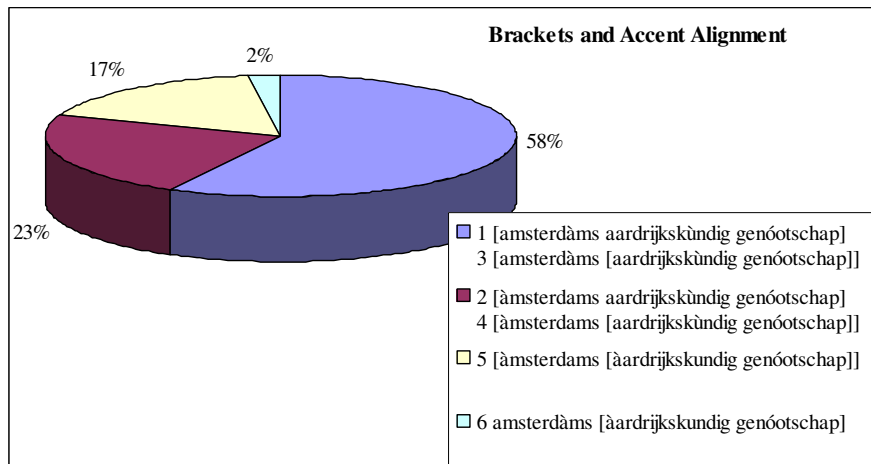
The constraints used in Table 28 are WRAP-XP, OUTPUT-OUTPUT CORRESPONDENCE (O-O CORR), LEFTALIGN, and ALIGN-XP. The constraint WRAP-XP demands that the elements of an input morphosyntactic constituent of type XP must be contained within a Phonological Phrase in the output representation (Truckenbrodt 1999). WRAP-XP also has the effect of deleting the phrase boundaries after each maximal XP (Horne et al. 1999: 73). O-O CORR, as we described in Chapter 3, demands that all variants of the same individual word parts have the same elements and prosodic structure. LEFTALIGN aligns an accent with the left edge of a phonological phrase, and ALIGN-XP is a constraint which aligns syntactic and prosodic boundaries (Selkirk 2000). Horne et al. (1999) assume a similar analysis. They observe:

“The ALIGN-XP constraint makes each maximal phrase category a prosodic phrase, and can come into conflict with WRAP-XP, because the latter constraint has the function of ‘wrapping’ an entire utterance into one prosodic phrase, whereas ALIGN-XP, unconstrained, would lead to an utterance having as many prosodic phrases as there are maximal phrase categories.” (Horne et al. 1999: 73).

This is exactly the way we analyse our data. The most important requirement for getting prosodically recursive phrases is the interaction between a constraint that ‘places the brackets’, i.e. that provides the right prosodic structure, and a constraint that aligns accents to these brackets. This is realized by WRAP-XP, ALIGN-XP and LEFTALIGN in Table 28. In Figure 56 we see that the percentages are as we would predict, following the constraint ranking in Table 28. On the other hand, all possible accent structures occur, including the less optimal ones.

Table 28 OT ranking (the output numbers refer to the numbers in Figure 56)

	WRAP-XP	O-O CORR	LEFTALIGN	ALIGN-XP
1			*	*
2		*!		*
☞ 3			**	
4		*!	*	
5		*!*		
6	*!			*

Figure 56 Observed percentages of different accent structures for the syntactically recursive NPs<sup>27</sup>

Output 3, [amsterdàms [aardrijkskùndig genóotschap]], is the optimal realization of the ranking in Table 28. It has the same realization as output 1 with regard to the accentuation pattern, which

<sup>27</sup> The percentages in Figure 56 are somewhat different from the percentages mentioned in the text. This discrepancy is due to the fact that for this graph we did not count the items with non-flexible accent patterns in Adjectives 1. We were most concerned with the patterns of Adjectives 2. Items with double accents are left out here as well for the sake of clarity.

indeed occurs in a vast majority of the cases. However, output 3 has an (inaudible) recursive phrase structure, while output 1 has a flat phrase structure. In Simulated Annealing, Outputs 2/4 and 5 can probably be variants if they form local optima. Coetzee's model (2004) would also predict the variants as suboptimal alternatives, with the cut-off point, based on the observations, set between WRAP-XP and O-O CORR. In a Stochastic model they could also be variants if LEFTALIGN and ALIGN-XP overlap with O-O CORR, in combination with a high noise level. Output 5 is the only candidate with an audible recursive phrase structure, which is due to the early accents. Output 6 occurs only in 2% of the cases and can therefore be seen as noise. It violates the highest-ranked constraint WRAP-XP.

The outputs 1 and 3, and 2 and 4, are not distinguishable auditorily. Our OT analysis happens to prefer recursive structures, however. Although our data show that a recursive phrase structure with an aligned accent pattern is not the preferred option, the vast majority of the data possibly does have a recursive phrase structure. If our analysis is right, this implies another argument against the Strict Layer Hypothesis.

Simultaneously with our analysis, Gussenhoven (2005) independently proposes a recursive analysis of multiply premodified NPs in English as procliticized phrases. He claims that the distribution of accents is determined by the interaction of lexical and post-lexical accent rules, with two lexical levels and one post-lexical level. The outputs of Level 1, where accents on prenuclear and nuclear feet are assigned and where Level 1 affixation is arranged, are input to the rules of Level 2, such as the COMPOUND RULE and Level 2 affixation. Level 2 outputs in their turn are input to a postlexical level, where phrases are formed. At the postlexical level Alignment constraints regulate the phrasal accent distributions. Differences in accent structures are, in this account, caused by the different accent distributions that leave the lexicon.

Gussenhoven claims that cyclic derivation, with OUTPUT-OUTPUT CORRESPONDENCE, does not work, because such a derivation would wrongly predict accentuation at the right boundary (output from Level 2) between an embedding premodifier and the head, as on *-teen* in *[[FIFTEEN] [Japanese conSTRUCtions]]* (capitals following Gussenhoven). Essential here is the constraint ALIGN( $\emptyset$ , T\*, RT)

(AlignRight), which aligns accents to the right edge of phonological phrases. The presence of the right boundary in combination with  $\text{ALIGN}(\emptyset, T^*, \text{RT})$  causes cyclicity to predict the wrong accent pattern. The solution Gussenhoven proposes is that the premodifiers are clitics, with a lefthand  $\emptyset$ -edge without at the same time creating a righthand  $\emptyset$ -edge. This means that *fifteen* does not embed *Japanese constructions*, but extends it. That is why he calls these phrases ‘procliticized phrases’. The absence of the right edge deprives  $\text{ALIGN}(\emptyset, T^*, \text{RT})$  of its erroneous influence. Besides the alignment constraints, the constraint  $\text{NOREMOTECLASH}$ , a constraint that demands that accents are more than one unaccented syllable apart, is needed (firstly for this data) to avoid *-nese* in *Japanese* being accented. In fact, this constraint bans alternation, which makes it an odd constraint with respect to the universality of OT constraints.

If we apply this analysis to our data, we come up with output 5 (*cf.* Figure 56) as the optimal candidate, while our more frequently observed outputs 1/3 and 2/4 crash on  $\text{ALIGN}(\emptyset, T^*, \text{Left})$ . Candidate 1 in Table 29, which corresponds to our winning candidates 3 and 1 in Figure 56, would be the worst candidate in Gussenhoven’s account. We illustrate this in Table 29a,b.

Table 29a Output candidates of our data using Gussenhoven’s account, input is output from Level 2

Input:	AMsterDAMS	AARDrijksKUNdig	geNOOTschap
Candidates:			
1	[amsterDAMS	[aardrijksKUNdig	geNOOTschap]
2	[AMsterdams	[aardrijksKUNdig	geNOOTschap]
3	[AMsterdams	[AARDrijkskundig	geNOOTschap]
4	[amsterDAMS	[AARDrijkskundig	geNOOTschap]
5	[AMsterDAMS	[AARDrijksKUNdig	geNOOTschap]
6	[AMsterDAMS	[AARDrijkskundig	geNOOTschap]
7	[AMsterdams	[AARDrijksKUNdig	geNOOTschap]
etc.			

Table 29b OT analysis of our data using Gussenhoven's account

	Dep-IO(Accent)	ALIGN( $\phi$ , T*, R)	ALIGN( $\phi$ , T*, Left)	NOCLASH	NOREMOTECLASH	MAX-IO (Accent)
1			*!			**
2			*!			**
☞ 3						**
4			*!	*	*	**
5				*!	***	
6				*!	**	*
7					*!	*
etc.						

Output candidates 1, 2, and 4 violate ALIGN( $\phi$ , T\*, Left) because they have no accent on either *am-* or *aard-*. This is already fatal for these candidates, and candidate 1 violates it even twice. NOCLASH is violated by candidates 4, 5, and 6, because two adjacent syllables are accented. Candidate 4 and the candidates with multiple accents per word (5, 6, 7, etc.), fail on NOREMOTECLASH, because only one unaccented syllable intervenes between the accented ones in one or more places. Notice that candidate 7, and probably more of such multiply accented phrases, falls out later than candidates 1 and 2, while these are observed in our data, and they are the preferred outputs according to our analysis in Table 28.

What is more, standard OT only allows one derived level, namely the output. Revising OT to a version in which different levels each have their own constraint ranking (*cf.* Kiparsky 1982a,b, D.B. den Ouden 2004) would reduce the restrictiveness of the theory with a proliferation of the same kind of constraints that work at different levels. Furthermore, a more modular model would prohibit the interaction of different influences on outputs and consequently reduce the possibility of accounting for so-called ‘conspiracies’ of different

influences in phonology, which turned out to be one of the great merits of OT.

Gussenhoven proposed this analysis for English phrases. For Dutch, it appears not to predict the right results, and the same holds for French (Gussenhoven personal communication). Our results also contradict the findings of Visch (1989), who concludes that in Dutch the tendency to choose early accent placement over O-O CORR is much stronger than in English. Because our own analysis can cope with our Dutch data better, we will stick to our straightforward and simple analysis.

#### **4.5. Conclusion**

The results of this experiment show that the prosodic recursion hypothesis holds: recursion does exist in prosody. We have shown that phonological phrase boundaries are often signaled by early accent placement, though not always. Our data also show that prosodic structure is less linear than assumed in the Strict Layer Hypothesis and derived hypotheses. In 22% of our data, the second adjective in a syntactically recursive noun phrase was perceived as having an early accent, and therefore it is reasonable to assume that these phrases were prosodically recursive as well. What is more, the OT analysis suggests that possibly a vast majority of our data may have a recursive phrase structure, though this structure is inaudible without the early accents. The analysis has a clear preference for this structure. In other words, the results of our experiment must be understood as additional evidence for a more prominent place for recursion in phonology. It would have been peculiar, after all, if recursion were not to exist in prosody, while it is currently seen as the most characteristic feature of linguistic syntax (*cf.* Hauser et al. 2002), and it is found in all kinds of structures in the world, such as nature itself, visual art, and music, as we showed in section 4.2. One could say that syntax consists of tacit cognitive structure, while phonology deals with the cognitive structure of physical behavior, which sets it apart. But especially the finding that music, as another

physically performed cognitive behavior, shows recursive structure, is an argument for assuming recursion in phonology as well.

Yet the results also show that the embedded phonological phrases, e.g. *aardrijkskundig genootschap*, do not behave identically to the maximal phonological phrases, e.g. *Amsterdams aardrijkskundig genootschap*, in the sense that the maximal phonological phrases are early accented significantly more often. Recall that the OT analysis showed that these seemingly non-recursive structures preferably do have recursive phrase structure. The difference is that the early accent placement, as a signal for phrase structure, is a less strong tendency than assumed. Clearly, there is a lot of optionality involved, not only for the subjects to apply the process of early accent placement, but also with regard to the recursivity of prosodic structure itself. The results confirm the observation that there is no one-to-one mapping from syntax to prosody, because optionality in syntactic structure would not be possible.

Another conclusion which came out of our analyses is that the perception of an early accent is often based on the de-accenting of the main stress syllable (see also Horne 1990, Gussenhoven 1991). Apparently pitch accent perception is relative: the absence of an accent on the expected position is interpreted by the listener as an accent on the next accentable position to the left. This is another indication that listeners base their perception not only on the acoustic signal alone (*cf.* Chapter 3); some strategy for retrieving the linguistic structure of the utterance must also play a role.

The propensity to get an early pitch accent does not depend on the number of intermediate syllables, as was assumed in the Eurhythmy rules (Hayes 1984). It does depend on rhythmic timing, however. Again, we find the ideal interval of ~ 600 ms between accents. This finding confronts us with the question whether early accent placement should still be assumed as a phrase-marking device or, conversely, as a purely rhythmic phenomenon. Although the finding that rhythmic timing plays an important role in early accent placement seems to weaken the assumption that early accent placement is a phrasing phenomenon, we cannot reject this assumption. The fact that it only occurs in phrases, and not in the same adjectives as individual words, is still an important indication

that it is a phonological structuring device. Besides that, we showed that the Equal Spacing Constraint did not apply between early and non-shifted accents.

In sum, we can conclude that we found strong indications for recursion in phonology, on the basis of auditory as well as on the basis of acoustic pitch and duration data.