Chapter 7

Prosodic Realization

7.1 Introduction

This chapter describes ProsR, the module of our SLG system CONPAS (described in Chapter 5) that is responsible for the prosodic realization of utterances. It receives its input from SynR, the syntactic realization module (described in Chapter 6), in the form of syntactic trees augmented with semantic and informational features. Its task involves, in chronological order, prosodic phrasing, pitch accent placement, intonational phrasing, and tune choice. Its output consists of strings of words with prosodic annotations, which are passed on to the speech synthesiser.

At least two aspects of prosodic realization have already received sufficient attention in the previous chapters. Phonological phrasing, which is the most important part of prosodic phrasing for our purposes, was investigated in depth in Chapter 3. The φ-algorithm implemented in ProsR is virtually identical to the one developed there, and will therefore only be summarised in Section 7.2. Similarly, ProsR uses the $\mathcal{NEW}$ algorithm for intonational phrasing developed in Chapter 4, and a description of it will not be repeated here.

The focus of the current chapter is on pitch accent placement. Following our research objective of designing, building and testing a linguistically-based SLG system that generates pitch accents most acceptable to experts, we present a rule-based generative grammar for pitch accent placement. The point of departure is the preliminary grammar that was presented in Section 2.5. Further development is directed by the research questions formulated in Section 1.3: how adequate is the theory, and how can it be improved? In Section 7.3, we argue for a number of modifications and extensions, part of which are motivated by inadequacies encountered during the development of PTDGen and WFDGen. In particular, we will pay attention to the correct accentuation of complex numerals (Section 7.3.5). The analyses in the current chapter rely on exemplary data. A formal evaluation of the accent placement algorithm (and of that for intonational phrasing as well) will be described in the next chapter.

Another aspect of prosodic realization, tune choice, is strictly speaking beyond the scope of our research objective. However, since an SLG system cannot function properly without it, a simple, provisional implementation of tune choice will be outlined in Section 7.4.

Prosodic realization is implemented in GRAMTSY (Section 6.3.1) as a grammar of transformational rules inserting features into labelled-bracket structures. Following the practice of the previous chapters, we will avoid discussing the implementation details, and describe prosodic realization in ProsR at a higher level of abstraction.

7.2 Prosodic phrasing

Prosodic phrasing in ProsR concerns the derivation of the higher-level prosodic constituents: the phonological phrase (φ) and the intonational phrase (π).[1] Recall that the prime motivation for deriving φ's is that they serve as the domain of application for the Rhythm Rule (RR), an accent deletion rule that modifies the accent distribution (Section 2.5.3). In addition, prosodic constituents are used as the domain of application for segmental postlexical rules (Section 5.6.1), and their boundaries may trigger preboundary lengthening and pause (Section 5.6.2).

7.2.1 Phonological phrase

In Chapter 3, several algorithms for phonological phrasing were evaluated, and it was found that none of them was fully adequate, and an improved algorithm for φ-phrasing was proposed, which has been implemented in ProsR. Since this subject has been amply investigated in Chapter 3, we will confine ourselves to a recapitulation of the algorithm (1).

(1) **Phonological phrase construction for Dutch**

A φ-boundary occurs after every content word W unless

a. W is an adjacent modifier, or

b. W is an A or Adv directly preceding a verb.

The algorithm inserts a φ-boundary after every content word, which may be a noun[2], verb[3], adjective, adverb[4]. The class of content words is extended with numerals and letters; for instance, the numeral *een* 'one' and the letter *A* in *platform one A*. However, the content word can neither be an adjacent modifier (2), nor an adjective or adverb directly preceding a verb. The definition of adjacent modification has been extended to numerals, in order to account for their accentuation (which will be discussed in Section 7.3.5).

[1] The lower-level prosodic constituents are derived by SpeechS (Section 5.6.1).

[2] Recall from Section 2.6.3 that 'pronouns' are syntactically analysed as determiners.

[3] This excludes functional verbs, i.e. auxiliaries and copula.

[4] This also excludes adverbs that are traditionally analysed as pronominal adverbs, such as *er* 'there' and *daarvoor* 'therefore'.

(2) **Adjacent modification**

For every $X, Y \in \{N, A, Adv, Num\}$:
X is an *adjacent modifier* of Y if and only if

a. X is left-adjacent to Y, and
b. X is dominated by YP.

Compounds are opaque with respect to φ-phrasing, that is, (1) will never insert a φ-boundary after a content word that is a non-final constituent of a compound word. Examples of φ-phrasing are repeated in (3) and (4). Notice that *erg* is an adjacent modifier of *mooie*, which is in turn an adjacent modifier of *tuin*; *ijskoud* is an adjective directly preceding a verb.

(3) a. $(_{NP}$ een $(_{AP}$ $(_{Adv}$ erg $)$ $(_{A}$ mooie $)$ $)$ $(_{N}$ tuin $)$ $)$
a very beautiful garden

b. $[_\varphi$ een $\overset{*}{\text{e}}$rg m$\overset{\emptyset}{\text{o}}$oie t$\overset{*}{\text{u}}$in $]$

(4) a. bier moet $(_{VP}$ $(_{A}$ ijskoud $)$ $(_{Aux}$ worden $)$ $(_{V}$ gedronken $)$ $)$
beer should ice-cold be drunk
'serve beer cold as ice'

b. bier $[_\varphi$ moet $\overset{*}{\text{i}}$jsk$\overset{\emptyset}{\text{o}}$ud worden gedr$\overset{*}{\text{o}}$nken $]$

As mentioned at the end of Section 3.3.3, the algorithm still suffers from a problem with unparsed auxiliaries and clitics. In (5), the verb *neemt* /nemt/ and the pronoun *u* /y/ are subject to *resyllabification*, that is, they are pronounced as a word with two syllables: /nem-ty/. Since it is generally assumed that the domain for resyllabification is the prosodic word (ω), it follows that both words must belong to the one and the same ω. However, our algorithm predicts a φ-boundary between them. According to the principles of prosodic phonology, a ω cannot dominate a φ, and we thus end up with a contradiction. Moreover, the resulting lack of resyllabification in the synthesised version of (5) leads to a conspicuously unnatural pronunciation.

(5) $[_\varphi$ vervolgens $]$ $[_\varphi$ neemt $]$ $[_\varphi$ u de sneltrein $]$
next take you the fast-train

In fact, Dutch has a limited number of function words like *u* that behave like phonological *clitics*, which attach to an appropriate host to form a single ω(Booij 1995, ch. 8). Accordingly, ProsR contains a cliticization rule that moves a clitic to the preceding φ, which solves this problem. A similar rule moves auxiliaries at the end of a VP to the preceding φ (see Section 3.2.2 for a description of this problem). Admittedly, these are rather ad hoc rules, and recent approaches in Optimality Theory may allow a better solution (Truckenbrodt 1999; Post 2000; Selkirk 2000).

7.2.2 Intonational phrase (π)

In Section 2.3.7, we distinguished between two phonological interpretations of the 'theory-neutral' concept of an intonational phrase: (1) the *Intonational Phrase* (π) as a prosodic category in the theory of prosodic phonology, and (2) the *Association Domain* (AD) as an independent domain for realizing an intonation contour. Since the first interpretation was shown to give rise to inconsistencies, we adopted the second view. The π, although no longer associated with intonation, is nevertheless preserved, because it useful as the domain of application for postlexical rules, and because its boundaries would appear to trigger pause insertion and preboundary lengthening. ProsR uses the following mapping rule for π-phrasing, which is essentially the rule proposed in Nespor and Vogel (1986, ch. 7), the difference being that not only root sentences, but all sentences, including bound and relative, trigger an π-boundary in (6).

(6) **Intonational phrase construction**

A π-boundary occurs:

a. before and after an appositional string;
b. before and after a sentence.

7.3 Pitch accent placement

A linguistic solution to pitch accent placement – as aimed for in the current work – is a generative grammar that takes into account all relevant factors in order to produce a correct accent distribution. In Section 2.5.4, we outlined a concise grammar for this goal that required four types of information as input: the surface string, the phonological phrases, the argument structure, and the focus distribution. Given this input, the grammar itself consisted of only two rules. First, the Sentence Accent Assignment Rule (SAAR), which accounts for the focus-accent relation. Second, the Rhythm Rule (RR), which accounts for rhythmical adjustments to the pitch accent distribution. This grammar constitutes the core of the implementation of pitch accent placement in ProsR described below. It will be shown, however, that the formulation of SAAR needs to be somewhat refined. In addition, it will be argued that an extra rule called *NP Premodifier Deaccentuation* is necessary. Furthermore, a lexically-driven rule called the *Lexical Rhythm Rule* is added. Finally, it will be shown that with the help of just one extra rule, *Unit Rhythm Rule*, an adequate accentuation of Dutch numerals is obtained.

7.3.1 Sentence Accent Assignment Rule

The relation between focus and accent forms the point of departure for pitch accent placement. This is captured in the first part of SAAR (Gussenhoven 1992b) – repeated here as (7)– which requires every focused constituent to be accented. The second part makes an exception for predicates that meet the specified condition.

(7) **Sentence Accent Assignment Rule** (SAAR)

If focused, every predicate, argument, and modifier must be accented, with the exception of a predicate that, discounting unfocused constituents, is adjacent to an argument.

An example from WFD is given in (8-a), where focus is indicated by underlining. The focused predicate *gebruikt* remains unaccented, because its adjacent argument *als kleurstof* is focused. In (8-b), the same argument is not focused, but the predicate can remain unaccented by virtue of its focused argument *het sap van de bladeren*; notice that the intervening unfocused argument *daarvoor* is ignored. In contrast, intervention of the *focused* modifier *zelden* forces the predicate to be accented in (8-c).

(8) a. het sap van de bladeren wordt als kleurstof gebruikt
the juice of the leaves is as dye used
'the juice of the leaves is used as dye'
[$_{\text{ARG}}$ het s$\overset{*}{\text{a}}$p van de bl$\overset{*}{\text{a}}$deren] wordt [$_{\text{ARG}}$ als kle$\overset{*}{\text{u}}$rstof]
[$_{\text{PRED}}$ gebruikt]

b. het sap van de bladeren wordt daarvoor gebruikt
the juice of the leaves is for-that used
'the juice of the leaves is used for that purpose'
[$_{\text{ARG}}$ het s$\overset{*}{\text{a}}$p van de bl$\overset{*}{\text{a}}$deren] wordt [$_{\text{ARG}}$ daarvoor]
[$_{\text{PRED}}$ gebruikt]

c. het sap van de bladeren wordt zelden gebruikt
the juice of the leaves is rarely used
[$_{\text{ARG}}$ het s$\overset{*}{\text{a}}$p van de bl$\overset{*}{\text{a}}$deren] wordt [$_{\text{MOD}}$ z$\overset{*}{\text{e}}$lden]
[$_{\text{PRED}}$ gebr$\overset{*}{\text{u}}$ikt]

From a practical point of view, it turns out to be more convenient to implement SAAR in two steps: (1) accent all focused constituents; (2) deaccent predicates where appropriate.

7.3.1.1 Accenting focused constituents

Recall that the focus distribution is either explicitly specified by the document planner (in terms of the features `foc` and `back`) or determined by the default focus rules in SynR. The latter are based on the content-function word dichotomy, foregrounding content words and backgrounding function words (Section 6.3.2). The first step in ProsR is thus to accent every word that carries the `foc` feature, which amounts to adding the feature `acc` to it. There are two cases that deserve special attention: compounds capable of bearing more than one accent, and compound verbs.

Compounds A compound adjective like *roodbruin* 'red-brown' can have an accent on its first part or on both parts, depending on the context of the word. Therefore, it is not sufficient to attach the feature acc to the word *roodbruin*; ProsR must have access to the internal structure of the word, or more precisely, to those parts that can carry an accent.[5] This means that syntactic trees with orthographic words as terminals will not do as input to ProsR.

Yet, generating the morphological structure of every word would be an exaggeration, because most of it is irrelevant for the purposes of accentuation. To be precise, the syntactic structures generated by SynR must include the internal structure of compounds if and only if this is relevant for accent placement, i.e. for those words in which the place of the accent depends on the context or which can take multiple accents. This class of compounds is called *phrasal compounds* in Varga (1998), because with respect to their accentuation they are closer to a syntactic phrase than a word. For Dutch, this class includes many compound adjectives and adverbs, all compound verbs, most compound numerals, and some compound nouns.[6] Their internal structure is listed in SynR's lexicon, and consequently ends up in the syntactic structures passed on to ProsR. All other words that are compounds from a morphological point of view, however, have a fixed accent position and are treated by SynR and ProsR like ordinary non-compound words.

The accentuation of phrasal compounds initially depends on the focus distribution. If focused, however, the feature acc is not attached to the compound itself, but to one or more of its parts. In the case of adjectives, adverbs, and numerals, every part receives an accent. Within nouns, only those parts that represent content words receive an accent; not the function words (e.g. conjunctions or prepositions). For example, *aan* 'at' is not accented in *Zandvoort-aan-Zee* 'Zandvoort-at-sea' (the name of a city). Some of these initial accents will be subject to deletion by rhythmic rules later on.

Compound verbs The accentuation of focused compound verbs, which consist of a particle and a verbal part, requires special care. In contrast to the other phrasal compounds, compound verbs can carry only one accent, which always goes to the particle. However, either the verbal or the particle part may have been moved out of its original position for syntactic reasons. For

[5] One might consider introducing different features to indicate an accent on the first part, the second part, or both parts of a compound. These would then have to be translated to corresponding escape codes that instruct the speech synthesiser to realize the accent in the intended location. One of the drawbacks is that this requires the speech synthesiser to do a morphological analysis to detect the parts of the compound. Apart from the fact that KUNTTS is currently not capable of interpreting such escape codes, it is not a structural solution either, because some compounds have three or more members. Examples of this are *Zoeterm$\overset{*}{e}$er-Centrum-W$\overset{*}{e}$st* 'Zoetermeer-Centre-West' (name of a railway station), *kat-en-m$\overset{*}{u}$is-spel* 'cat-and-mouse-game', and *doe-het-z$\overset{*}{e}$lf-winkel* 'do-it-yourself-shop'. To account for all these possibilities, even more accent features and corresponding escape codes would be required.

[6] For practical reasons, we included compounds with an internal structure that is so complex that the speech synthesiser cannot determine the position of the accent correctly.

example, both parts remain together in (9-a), but the verbal part is fronted in (9-b), whereas the particle is moved leftwards in (9-c). With respect to (9-b), we have to make sure that the verb *stapt* is recognised as the verbal part of a compound, preventing accentuation as a normal content verb. Similarly, *over* must be identified as a particle (not a preposition) in (9-c), which is accented only if the compound verb that it originates from is focused.

(9) a. dat u daar ($_V$ ($_P$ over) ($_V$ stapt))
that you there `prt` transfer
'that you transfer there'

b. u ($_V$ stapt $_i$) daar ($_V$ ($_P$ over) e_i)
you transfer there `prt`
'that's where you transfer'

c. om daar ($_P$ over $_i$) te ($_V$ e_i ($_V$ stappen))
to there `prt` `prt` transfer
'to transfer there'

Recall from the previous chapter that determining the focus distribution precedes determining the word order, and that the feature `foc` is attached to a compound verb, not to its constituents. Therefore, if the verbal part is moved, as in (9-b), it will not carry the feature `foc` with it, and thus cannot be accented. This automatically leads to the correct prediction that the verbal part of a compound verb is indeed always unaccented. In order to account for the accentuation of the particle, SynR copies the feature `foc` to the particle *before* it is moved. As a result, it will be focused if and only if the compound verb is focused, and thus correctly accented, even when it is moved to another position, as in (9-c).[7]

To sum up so far, we have described the implementation of the first part of SAAR, which tells us to accent all focused material. We have seen that for normal words, the focus-accent relation is simply one-to-one, whereas for phrasal compounds, the relation is more complex, but can be satisfactorily implemented if SynR carries out the preliminary work.

7.3.1.2 Deaccenting the predicate

To continue with the second step of SAAR: a focused predicate is deaccented if it is adjacent to a focused argument, ignoring any intervening *unfocused* material. Since SynR marks predicates and arguments with the features `pred` and `arg` respectively, this is relatively straightforward to implement. For each focused predicate, ProsR checks if it has a focused argument to its left or right, and if no other focused constituent intervenes between these. If so, the feature `acc` is removed from the predicate.

However, there are a few special circumstances. The first has to do with the special properties of compound verbs. Second, SAAR is stated in terms

[7] As a matter of fact, not only `foc`, but also the predicate-marking feature `pred` is copied onto the particle, to facilitate the deaccentuation step of SAAR later on.

of focused *semantic constituents*, whereas CONPAS assumes that focus is distributed over *words*; this requires some mechanism of focus projection from words onto constituents. Third, there are a number of exceptions to SAAR, which require extra conditions. These issues will be discussed below.

Compound verbs Recall from the previous section that the `foc` feature of a compound verb is copied onto its particle, because if the verb is focused, it is the particle that must be accented, even when it is extracted from the verb. For similar reasons, the `pred` feature is moved from the compound verb onto its particle: SAAR must be able to deaccent the particle if it stands in the required relation to an argument, just as it would in the case of an ordinary verb. An example of this is (10), where the particle *over* of the focused compound verb *overstappen* is adjacent to its focused argument *op de stoptrein*, which allows the particle to remain unaccented.

(10) u hebt v$\overset{*}{\text{i}}$jf min$\overset{*}{\text{u}}$ten om op de ($_{N<foc,arg>}$ st$\overset{*}{\text{o}}$ptrein)
you have five minutes for on the slow-train
($_{P<foc,pred>}$ over$_i$) te ($_{V<foc>}$ e_i ($_V$ stappen))
`prt` `prt` transfer
'you have five minutes to transfer to the slow train'

Focus projection For SAAR to apply to a predicate, a focused argument must be adjacent. However, if a complex argument contains multiple content words of which some are focused and some are not, this raises the question what makes the argument as a whole count as focused. According to Gussenhoven (1992b, Appendix), the *head* of a phrase that functions as argument must be focused; it is not sufficient to have only a modifier focused. This explains why in the question-answer pair (11) the predicate *vertrokken* is focused but unaccented, whereas it must be accented in in (12).

(11) *Question:* waarom wachten we?
'why are we waiting?'

Answer: de tr$\overset{*}{\text{e}}$in is al vertrokken
the train is already gone

(12) *Question:* waarom wachten we op de trein naar Amsterdam?
'why are we waiting for the train to Amsterdam?'

Answer: de trein naar R$\overset{*}{\text{o}}$tterdam is al vertr$\overset{*}{\text{o}}$kken
the train to Rotterdam is already gone

There appear to be at least two ways to implement this. The first approach is to make SAAR (actually, the GRAMTSY rules that implement SAAR) scan the inner structure of an argument in order to determine if its head is focused.

However, since an argument can be an NP, PP, AP, or S with a complicated syntactic structure, even a coordinated structure, this would make the implementation of SAAR rather complex. The alternative approach is to assume that the fact that an argument is focused is given in the input to ProsR. This implies that it belongs to the tasks of SynR to project focus from the head of an argument onto its maximal projection. This approach is more attractive, as the distribution of (agreement) features along constituents is one of the tasks in syntactic realization anyway. Moreover, it allows for a concise formulation of SAAR in ProsR, abstracting away from syntactic details.

SynR implements this approach by means of a set of rules akin to those proposed by Selkirk (1984, 1995) in her focus projection theory. However, since Selkirk's generalised rules can be shown to make incorrect predictions in some cases (Gussenhoven 1992b; Gussenhoven 1999), we prefer the more restricted formulation below.

(13) **Focus Projection Rules**

a. If the head of an NP is focused, then the NP is focused as well.

b. If the head of an AP is focused, then the AP is focused as well.

c. If the NP argument of a P is focused, then the PP is focused as well.

d. If the main verb in an S is focused, then the S is focused as well.

e. If all the conjuncts of a coordination are focused, then the coordination is focused as well.[8]

Rule (13-a), for example, projects focus from the N *trein* onto the NP argument *de trein* in (11), allowing deaccentuation of the predicate *vertrokken*. In contrast, although rules (13-a) and (13-c) project focus from *Rotterdam* onto the PP *naar Rotterdam* in (12), there are no rules that license further projection. That is, the argument *de trein naar Rotterdam* remains unfocused, precluding deaccentuation of the predicate *vertrokken*.

Topicalised arguments According to Gussenhoven (1992b), an exception to SAAR is that topicalised arguments cannot cause the predicate to be deaccented. Consider the example in (14), where everything except *ik* is focused. Whereas the predicate *slaat* may remain unaccented in (14-a) (as predicted by SAAR), it must be accented in (14-b), even though the topicalised direct object *Jan* is adjacent to it.

[8] The set of conjuncts does not include the conjunction (e.g. *and*)

(14) a. Ja⃰n slaat me
John beats me

b. Ja⃰n sla⃰ ik
John beat I

For most types of text, topicalised arguments other than the subject appear to be rare; they are commonly assumed to be bound to certain types of discourse transitions (see e.g. Rats (1996)). Despite the fact that they rarely occurred in PTD or WFD, this restriction on SAAR is implemented in ProsR by checking that the argument to the left of the predicate is not topicalised (i.e. lacks the feature `top`).[9]

Extraposed arguments One of the observations resulting from the current work is that deaccentuation of the predicate often appears to be less appropriate with extraposed arguments. Consider the data in (15): whereas the predicate *overstappen* must remain unaccented by virtue of the adjacent argument *op de stoptrein naar Schiedam Centrum* in (15-a), it can very well be accented if the same argument occurs to its right as a result of extraposition (15-b). In fact, accentuation of the predicate is obligatory when followed by an intonational boundary (15-c), if it is not to be interpreted as unfocused.

(15) a. u hebt a⃰cht minu⃰ten # om op de sto⃰ptrein naar
you have eight minutes for to the slow-train to
Schieda⃰m Ce⃰ntrum over te stappen
Schiedam Centre `prt` to transfer
'you have eight minutes to transfer to the slow train to Schiedam Centre'

b. u hebt a⃰cht minu⃰ten # om o⃰ver te stappen op de
you have eight minutes for `prt` to transfer to the
sto⃰ptrein naar Schieda⃰m Ce⃰ntrum
slow-train to Schiedam Centre

c. u hebt a⃰cht minu⃰ten # om ru⃰stig o⃰ver te stappen #
you have eight minutes for quietly `prt` to transfer
op de sto⃰ptrein naar Schieda⃰m Ce⃰ntrum
to the slow-train to Schiedam Centre

Example (15-c) thus suggests an interaction between pitch accent placement and intonational phrasing, challenging our assumption that pitch accent placement precedes intonational phrasing. We will return to this issue in Section 7.3.6.4.

[9] In our syntactic analysis, subjects can also be topicalised. However, this exception to SAAR does not apply to topicalised subjects.

In sum, the second part of SAAR (deaccenting of the predicate) requires focus projection rules, because we need to know when an argument counts as focused. The implementation of deaccenting a compound verb requires some extra precautions, as the particle may be separated from the verbal part. In addition, deaccentuation is not allowed if a topicalised argument other than the subject is involved, and is blocked by an intervening AD-boundary if an extraposed argument is involved.

7.3.2 Deaccentuation of NP premodifiers

In addition to the deaccentuation of predicates by SAAR, there appears to be at least one other type of deaccentuation. Examples of this type occur in the NP's below.

(16)	a.	station N$\overset{*}{\text{i}}$jmegen station Nijmegen 'Nijmegen station'
	b.	minister-president K$\overset{*}{\text{o}}$k prime-minister Kok
	c.	de wedstrijd $\overset{*}{\text{A}}$jax-Fe$\overset{*}{\text{i}}$jenoord the game Ajax-Feijenoord
	d.	perron v$\overset{*}{\text{i}}$jf platform five
	e.	bladzijde v$\overset{*}{\text{i}}$erend$\overset{*}{\text{e}}$rtig, regel twa$\overset{*}{\text{a}}$lf page 34 line 12
	f	het woord 'tr$\overset{*}{\text{e}}$in' the word train
	g	een glas w$\overset{*}{\text{i}}$jn a glass wine
	h	meneer de v$\overset{*}{\text{o}}$orzitter mister the chairman

All of these involve a sequence of two closely related nominal parts of which the first always has a common noun. The point is that although all these examples can be interpreted as fully focused, only the second part is accented, whereas the first remains unaccented (even though it is a content word).

The syntactic structure of these examples is unclear. The structuralistic tradition suggests they constitute 'double-headed structures' (Haeseryn et al. 1997, p. 875-878). We will assume they form two NP's of which the first one serves as a premodifier of the second, akin to an AP modifying an N. That is, given the abstract phrase specification in (17-a), SynR will generate the syntactic structure in (17-b).

(17) a. `(Nijmegen<prop>(station<nonref>))`

b. $({}_{NP}\ ({}_{NP}\ ({}_{N}$ station)) $({}_{N}$ Nijmegen))

It wil be clear that neither SAAR nor RR can explain the lack of an accent on the NP premodifier: the former is not applicable, as there is no obvious predicate-argument structure present; the latter needs at least three accents to delete one of them. This forces the introduction of an additional deaccentuation rule:

(18) **NP Premodifier Deaccentuation** (NPPD)

Delete the pitch accents on an NP that serves as a premodifier of another NP.

7.3.3 Rhythm Rule

After SAAR and NPPD have applied, RR (19) deletes the intermediary accents in phonological phrases containing three or more accents.[10] The net effect of exhaustive application is that only the outermost accents in a phonological phrase are preserved. In order to prevent incorrect accent deletions, it is crucial that the boundaries of phonological phrases are correctly determined. Since RR and phonological phrasing for Dutch have been extensively discussed and tested in Chapter 3, we will not elaborate on this topic here.

(19) **Rhythm Rule** (RR)

$$* \rightarrow \emptyset \ / \ [_{\varphi} \ldots [_{\omega} *] \ldots [_{\omega} \_] \ldots [_{\omega} *] \ldots]$$

7.3.4 Lexical rhythm

There is a very limited class of quantifiers and adverbs for which the location of the pitch accent depends on the context (Gussenhoven 1983a). Examples of this are *allemaal* 'all of them', *altijd* 'always', and *meestal* 'often'. The pitch accent goes to the first part of the word if another accent follows in the same utterance (20-a), but goes to the second part if it is the last accent in the utterance (20-b).

(20) a. die trein rijdt $\overset{*}{\text{a}}$ltijd om z$\overset{*}{\text{e}}$ven $\overset{*}{\text{u}}$ur
that train drives always at seven hour

b. die trein rijdt alt$\overset{*}{\text{i}}$jd om zeven uur

[10]The notation is somewhat different from that in Section 2.5.3 – more precise actually – in order to facilitate the formulation of two additional rhythm rules below.

It is assumed that words belonging to this class are all phrasal compounds (cf. Section 7.3.1.1), which means that, if focused, they initially receive two accents. In addition, it is assumed that they are lexically marked with the feature `lrr`, originating from SynR's lexicon. This feature is then recognised by a special rule for lexical rhythm that deletes either the first or the second accent.

(21) **Lexical Rhythm Rule** (LRR)

$$* \;\rightarrow\; \emptyset \;\;/\;\; \left\{ \begin{array}{l} [_{\mathrm{U}} \cdots [_{\omega} \, [_{\omega} \; * \;] \; [_{\omega} \; \_ \;] \;]_{\mathtt{lrr}} \cdots [_{\omega} \; * \;] \; \cdots] \\ [_{\omega} \, [_{\omega} \; \_ \;] \; [_{\omega} \; * \;] \;]_{\mathtt{lrr}} \end{array} \right\}$$

That is, for every word marked with the feature `lrr`, if the remainder of the utterance contains at least one accent, then the accent on its second constituent is deleted, otherwise the accent on its first constituent is deleted. Compared with RR, there are two differences: (1) the left context of the word is irrelevant; (2) the right side context is not limited to a φ, but extends to the end of the utterance.

7.3.5 Accentuation of Numerals

Numerals occur in PTD during the generation of times and platform numbers, and in WFD in expressions of length or height. As a matter of fact, numerals are likely to occur in almost any application domain for SLG. As a consequence, the accentuation of numerals is a topic that deserves to be addressed. Ideally, their accentuation should be covered by the rules proposed so far. As will explained below, however, an additional deaccentuation rule is needed to obtain the intended results.

A numeral qualifies as a content word, and will therefore by default be focused, and thus initially accented. Numbers from 21 to 29, 31 to 39, and so on up to 91 to 99, are numeral compounds, which may carry two accents (22-a), just like any other phrasal compound (Section 7.3.1.1). Whenever they serve as an adjacent modifier, they are included into a φ with the following head, and subject to RR (22-b).[11]

(22) a. het zijn er z$\overset{*}{\text{e}}$sen n$\overset{*}{\text{e}}$gentig
it are there six-and ninety
'there are 96 of them'

b. z$\overset{*}{\text{e}}$sen n$\overset{\emptyset}{\text{e}}$gentig k$\overset{*}{\text{a}}$sten
six-and ninety cupboards
'96 cupboards'

In Section 6.2.3.4, we argued that for larger numeral phrases as in (23), *zesen negentig* should syntactically be analysed as a premodifier of *honderd/duizend*,

[11] All examples in this section are assumed to be fully focused.

analogously to (22-b). As a consequence, it becomes an adjacent modifier of the following numeral, included in the same φ, and subject to RR. This elegantly explains the accent deletions in (23).

(23) a. z$\overset{*}{\text{e}}$sen n$\overset{\emptyset}{\text{e}}$gentig h$\overset{*}{\text{o}}$nderd
six-and ninety hundred
9,600

b. tw$\overset{*}{\text{e}}$een d$\overset{\emptyset}{\text{e}}$rtig d$\overset{*}{\text{u}}$izend
two-and thirty thousand
32,000

c. tw$\overset{*}{\text{e}}$e h$\overset{\emptyset}{\text{o}}$nderd k$\overset{*}{\text{a}}$sten
two hundred cupboard

While this successfully explains the accentuation of most numerals, certain types of numerals appear to pose a problem. Consider the data in (24). At first sight, it seems as if the lack of an accent on *honderd* is again due to RR, which would imply that the whole phrase constitute a single φ.

(24) tw$\overset{*}{\text{e}}$e h$\overset{\emptyset}{\text{o}}$nderd (en) z$\overset{*}{\text{e}}$s
two hundred (and) six

However, this analysis cannot be right. Consider example (25-a). From a comparison to (23-c), we can infer that there is no φ-boundary after *twee*. On the basis of its similarity to (24), we are forced to assume that there is no φ-boundary after *honderd* either. As shown in (22-b), there is also no φ-boundary after *zesen*. Therefore, the numeral phrase in (25-a) must be a single φ. As a result, RR deletes the accents on *honderd* and *zes*. However, this gives to an unintended interpretation: without an accent, *zes* can only be interpreted as unfocused.

(25) *tw$\overset{*}{\text{e}}$e h$\overset{\emptyset}{\text{o}}$nderd (en) z$\overset{\emptyset}{\text{e}}$sen n$\overset{*}{\text{e}}$gentig
two hundred (and) six-and ninety
'296'

Likewise, the analysis of (24) as a single φ would imply (26) is a single φ as well (cf. (23-b), again resulting in an undesired deletion of the accent on *zes*. In fact, a unlimited number of counterexamples of this type can be constructed.

(26) *tw$\overset{*}{\text{e}}$een d$\overset{\emptyset}{\text{e}}$rtig d$\overset{\emptyset}{\text{u}}$izend (en) z$\overset{\emptyset}{\text{e}}$sen n$\overset{*}{\text{e}}$gentig
two-and thirty thousand (and) six-and ninety
'32,096'

In addition, *zes* cannot be deaccented in (27), where the whole numeral phrase serves as an adjacent modifier to the noun *kasten*.

(27) *tw$\overset{*}{e}$e h$\overset{ø}{o}$nderd (en) z$\overset{ø}{e}$sen n$\overset{ø}{e}$gentig k$\overset{*}{a}$sten
two hundred (and) six-and ninety cupboards
'296 cupboards'

Clearly, there must be a φ-boundary after *honderd* and *duizend* to prevent these undesired accent deletions. Actually, this is exactly what our φ-phrasing algorithm predicts. According to syntactic analysis of numerals proposed in Section 6.2.3.4, the last three examples all contain a coordination of two numeral phrases, of which the first is headed by *honderd/duizend*. See, for example, the syntactic structure of (25) in Figure 7.1. This syntactic analysis is supported by the fact that we can optionally insert the conjunction *en* 'and' after *honderd* in (25). Given this analysis, both *honderd* and the compound *zesennegentig* trigger a φ-boundary; see the predicted φ-phrases in (28).[12]

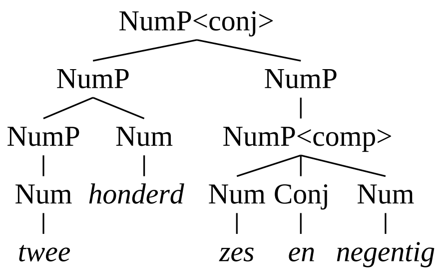

Figure 7.1: Syntactic structure of example (25).

(28) [$_\varphi$ tw$\overset{*}{e}$e h$\overset{*}{o}$nderd] [$_\varphi$ (en) z$\overset{*}{e}$s en n$\overset{*}{e}$gentig]

With respect to (27), the φ-phrases according to our algorithm are shown in (29). The accent on *negentig* is still deleted by RR, which is a correct prediction.

(29) [$_\varphi$ tw$\overset{*}{e}$e h$\overset{*}{o}$nderd] [$_\varphi$ (en) z$\overset{*}{e}$sen n$\overset{ø}{e}$gentig k$\overset{*}{a}$sten]

To sum up, we have argued that there must be a φ-boundary after *honderd/duizend*, which is in accordance with the predictions of our φ-algorithm. However, returning to the example we started with (24), we now face the problem of explaining why the accent on *honderd* is deleted, as this cannot be attributed to RR.

[12] Notice that *twee*, although it is a content word, is an adjacent modifier of *honderd*, and therefore incapable of triggering a φ-boundary. *Zes* is part of a numeral compound, and therefore invisible for the φ-phrasing algorithm.

Gussenhoven (1991a) addresses the same problem for English. On his account, the syntactic structure of the phrase *hundred and two thousand men*, which is left-branching (30-a), is mapped onto a right-branching prosodic structure with recursion at the φ-level (30-b). It is assumed that the application of RR is optional and cyclic, starting with the lowest φ. Consequently, the accent on *thousand* is deleted first, while that on *two* may be preserved.

(30) a. (((hundred and two) thousand) men)

b. $[_{\varphi''}$ h$\overset{*}{\text{u}}$ndred and $[_{\varphi'}$ $[_{\varphi}$ tw$\overset{*}{\text{o}}$th$\overset{\varnothing}{\text{o}}$usand] m$\overset{*}{\text{e}}$n]]

Although this solution would solve the above cases, it has a few drawbacks. First, the reason for the reversed mapping of syntactic onto prosodic structure remains somewhat mysterious. Second, except for a general tendency towards eurhythmy, it is unclear why RR should not delete the accent on *two* as well when it is applied to the φ''. Third, it requires the prosodic structure to be recursive, violating the Strict Layer Hypothesis, which we tried to adhere to in the current work (Section 2.2). A more serious problem, however, is that this approach cannot account for the lack of an accent on *honderd* in the phrases below.

(31) a. h$\overset{\varnothing}{\text{o}}$nderd v$\overset{*}{\text{i}}$eren d$\overset{*}{\text{e}}$rtig
hundred four and thirty
134

b. h$\overset{\varnothing}{\text{o}}$nderd v$\overset{*}{\text{i}}$eren d$\overset{\varnothing}{\text{e}}$rtig d$\overset{*}{\text{u}}$izend
hundred four and thirty thousand
134,000

c. h$\overset{\varnothing}{\text{o}}$nderd $\overset{*}{\text{e}}$en bl$\overset{*}{\text{a}}$dzijden
hundred one pages
101 pages

In these examples, the *first* accent in the utterance is lacking without affecting a broad focus interpretation. Since RR requires at least one accent in its left-hand context, it is by definition incapable of deleting the first accent in an utterance.[13]

By now, we might suspect that the reason that *honderd* is unaccented is simply because it virtually always unfocused. However, as (32) shows, there are a number of instances where *honderd* does receive an accent.

(32) a. h$\overset{*}{\text{o}}$nderd
hundred

b. n$\overset{*}{\text{e}}$gen h$\overset{*}{\text{o}}$nderd
nine hundred

[13] Similar problems arise in English: *a hundred and* $\overset{*}{\textit{o}}$*ne* p$\overset{*}{\textit{a}}$*ges*

c. n$\overset{*}{e}$genen dertig h$\overset{*}{o}$nderd
nine thirty hundred
39,000

d. h$\overset{*}{o}$nderd d$\overset{*}{u}$izend
hundred thousand
100,000

e. h$\overset{*}{o}$nderd k$\overset{*}{a}$sten
hundred cupboards

In conclusion, deletion of the accent on *honderd* cannot be accounted for by RR, at least not in the case of initial deaccentuation, nor can it be attributed to a lack of focus. This motivates the introduction of additional accent deletion rule. In the remainder of this section, we will formulate this rule, and show that its use is not limited to the case of *honderd*, but includes a number of similar phenomena.

The numeral *honderd* differs from the basic numerals, i.e. *een* up to *twintig* (1-20) and *twintig* up to *negentig* (20,30,...90), in the sense that it may function as a unit of measure. That is, (33) can paraphrased as *twee van honderd, en zes* 'two of hundred, and six', or *twee hondertallen en zes* 'two hundreds, and six'.

(33) tw$\overset{*}{e}$e h$\overset{\emptyset}{o}$nderd (en) z$\overset{*}{e}$s
two hundred (and) six

The same pattern is found for *duizend* 'thousand' and *miljoen* 'million' (34), and for other abstract units of measure when combined with a numeral (35).

(34) a. het zijn er tw$\overset{*}{e}$e d$\overset{\emptyset}{u}$izend $\overset{*}{e}$enen tw$\overset{*}{i}$ntig
it are there two thousand one-and twenty
'there are 2,021 (of them)'

b. het zijn er tw$\overset{*}{e}$e milj$\overset{\emptyset}{o}$en $\overset{*}{e}$enen tw$\overset{*}{i}$ntig
it are there two million one-and twenty
'there are 2,000,021 (of them)'

(35) a. tw$\overset{*}{e}$e $\overset{\emptyset}{u}$ur z$\overset{*}{e}$sen tw$\overset{*}{i}$ntig
two hour six-and twenty
'2:26'

b. tw$\overset{*}{e}$e m$\overset{\emptyset}{e}$ter v$\overset{*}{i}$jftig
two meter fifty
'two meters and fifty centimetres'

c. z$\overset{*}{e}$ven g$\overset{\emptyset}{u}$lden v$\overset{*}{i}$jfen n$\overset{*}{e}$gentig
seven guilder five-and ninety
'Fl 7.95'

d. $\overset{*}{e}$lf h$\overset{\emptyset}{o}$nderd na Chr$\overset{*}{i}$stus
eleven hundred after Christ
'1100 A.D.'

e. t$\overset{*}{i}$en gr$\overset{\emptyset}{a}$den onder n$\overset{*}{u}$l
ten degrees below zero

f. t$\overset{*}{i}$en m$\overset{\emptyset}{i}$nuten na $\overset{*}{a}$ankomst
ten minutes after arrival

Again, these deaccentuations cannot be accounted for by RR, assuming these phrases form single φ's, because this fails to explain why *een* in (34), *zesen* in (35-a), and *vijfen* in (35-c) must remain accented. As a matter of fact, our φ-algorithm predicts φ-boundaries after *duizend*, *uur*, and *gulden*.

However, it is not only the lexical form of these nouns that counts. As the examples in (36) show, deaccentuation is impossible if they are not used as units of measure. For the same reason, *honderd* must keep its accent in the last example of (32).

(36) a. ?de koers $\underline{\text{van de g}\overset{\emptyset}{\text{u}}\text{lden d}\overset{*}{\text{a}}\text{alde}}$
the price of the guilder dropped

b. ?het $\overset{\emptyset}{u}$ur van de w$\overset{*}{a}$arheid is $\overset{*}{a}$angebroken
the hour of the truth has come
'the moment of truth is upon us'

In anticipation of a more thorough investigation, we will assume that words used in this way are marked by the feature `urr`, which originates either from the input or from SynR. The feature triggers (37), which requires that the word so marked is deaccented if it is followed by another accent. *Honderd* cannot be deaccented if it is utterance-final, as is clear from the first three examples in (32), and neither can any other unit without causing a contrastive interpretation.

(37) **Unit Rhythm Rule** (URR)

$$* \rightarrow \emptyset \ / \ [_{\mathrm{U}} \cdots [_{\omega} \_\_]_{\mathbf{urr}} \cdots [_{\omega} *] \cdots]$$

However, there is one complication: the rule is blocked by an intervening AD boundary. In all of the following examples, deleting the accent on the unit preceding the AD-boundary results in a contrastive reading.

(38) a. [$_{AD}$ het zijn er z$\overset{*}{\text{e}}$ven honderd tw$\overset{*}{\text{e}}$een d$\overset{\emptyset}{\text{e}}$rtig d$\overset{*}{\text{u}}$izend]
it are there seven hundred two-and thirty thousand
[$_{AD}$ n$\overset{*}{\text{e}}$gen h$\overset{\emptyset}{\text{o}}$nderd $\overset{*}{\text{e}}$enen tw$\overset{*}{\text{i}}$ntig]
nine hundred one-and twenty
'there are 732,921 (of them)'

b. [$_{AD}$ tw$\overset{*}{\text{e}}$e honderd t$\overset{*}{\text{a}}$chtig m$\overset{*}{\text{e}}$ter] [$_{AD}$ en v$\overset{*}{\text{i}}$jfen vijftig
two hunderd eighty meter and five-and fifty
c$\overset{*}{\text{e}}$ntimeter]
centimetre

c. het huidige wereldrecord staat op [$_{AD}$ tw$\overset{*}{\text{e}}$e $\overset{*}{\text{u}}$ur] [$_{AD}$ z$\overset{*}{\text{e}}$s
the current world-record stands on two hours six
min$\overset{*}{\text{u}}$ten] [$_{AD}$ en tw$\overset{*}{\text{a}}$alf sec$\overset{*}{\text{o}}$nden]
minutes and twelve seconds

Since we assume that pitch accent placement *precedes* AD-derivation, AD-boundaries are not yet known when URR applies. Therefore, this condition cannot be included in the structural description of URR. We return to this problem in Section 7.3.6.4.

A final point that needs to be addressed is the mutual order of RR and URR. If URR precedes RR, URR can destroy the context for RR, and therefore prevent the deletion of accents. An example of this is (39). In contrast, changing the order, i.e. RR before URR, yields the correct output, as is shown in (40).

(39) Input: twee en dertig honderd$_{\texttt{ur}}$ twaalf
two and thirty hundred twelve
3,212

IA: tw$\overset{*}{\text{e}}$e en d$\overset{*}{\text{e}}$rtig h$\overset{*}{\text{o}}$nderd$_{\texttt{urr}}$ tw$\overset{*}{\text{a}}$alf

φ's: [$_{\varphi}$ tw$\overset{*}{\text{e}}$e en d$\overset{*}{\text{e}}$rtig h$\overset{*}{\text{o}}$nderd$_{\texttt{urr}}$] [$_{\varphi}$ tw$\overset{*}{\text{a}}$alf]

URR: [$_{\varphi}$ tw$\overset{*}{\text{e}}$e en d$\overset{*}{\text{e}}$rtig h$\overset{\emptyset}{\text{o}}$nderd$_{\texttt{urr}}$] [$_{\varphi}$ tw$\overset{*}{\text{a}}$alf]

RR: n/a

Output: ? tw$\overset{*}{\text{e}}$e en d$\overset{*}{\text{e}}$rtig honderd tw$\overset{*}{\text{a}}$alf

(40) Input: twee en dertig honderd$_{\texttt{urr}}$ twaalf

IA: twe$\overset{*}{\text{e}}$ en d$\overset{*}{\text{e}}$rtig h$\overset{*}{\text{o}}$nderd$_{\texttt{urr}}$ twa$\overset{*}{\text{a}}$lf

φ's: [$_\varphi$ twe$\overset{*}{\text{e}}$ en d$\overset{*}{\text{e}}$rtig h$\overset{*}{\text{o}}$nderd$_{\texttt{urr}}$] [$_\varphi$ twa$\overset{*}{\text{a}}$lf]

RR: [$_\varphi$ twe$\overset{*}{\text{e}}$ en d$\overset{\o}{\text{e}}$rtig h$\overset{*}{\text{o}}$nderd$_{\texttt{urr}}$] [$_\varphi$ twa$\overset{*}{\text{a}}$lf]

URR: [$_\varphi$ twe$\overset{*}{\text{e}}$ en dertig h$\overset{\o}{\text{o}}$nderd$_{\texttt{urr}}$] [$_\varphi$ twa$\overset{*}{\text{a}}$lf]

Output: twe$\overset{*}{\text{e}}$ en dertig honderd twa$\overset{*}{\text{a}}$lf

We finish this section on numerals with a more complex example (41), which shows that the combined operation of φ-phrasing, RR, and URR can account for accent distributions that are far from trivial.

(41) Input: honderd$_{\texttt{urr}}$ vier en vijftig duizend$_{\texttt{urr}}$ zeven
hundred four and fifty thousand seven
honderd$_{\texttt{urr}}$ tween en dertig kasten
hundred two and thirty cupboards
154,732 cupboards

IA: h$\overset{*}{\text{o}}$nderd$_{\texttt{urr}}$ v$\overset{*}{\text{i}}$eren v$\overset{*}{\text{i}}$jftig d$\overset{*}{\text{u}}$izend$_{\texttt{urr}}$
z$\overset{*}{\text{e}}$ven h$\overset{*}{\text{o}}$nderd$_{\texttt{urr}}$ twe$\overset{*}{\text{e}}$ en d$\overset{*}{\text{e}}$rtig k$\overset{*}{\text{a}}$sten

φ's: [$_\varphi$ h$\overset{*}{\text{o}}$nderd$_{\texttt{urr}}$]
[$_\varphi$ v$\overset{*}{\text{i}}$eren v$\overset{*}{\text{i}}$jftig d$\overset{*}{\text{u}}$izend$_{\texttt{urr}}$]
[$_\varphi$ z$\overset{*}{\text{e}}$ven h$\overset{*}{\text{o}}$nderd$_{\texttt{urr}}$]
[$_\varphi$ twe$\overset{*}{\text{e}}$ en d$\overset{*}{\text{e}}$rtig k$\overset{*}{\text{a}}$sten]

RR: [$_\varphi$ h$\overset{*}{\text{o}}$nderd$_{\texttt{urr}}$]
[$_\varphi$ v$\overset{*}{\text{i}}$eren v$\overset{\o}{\text{i}}$jftig d$\overset{*}{\text{u}}$izend$_{\texttt{urr}}$]
[$_\varphi$ z$\overset{*}{\text{e}}$ven h$\overset{*}{\text{o}}$nderd$_{\texttt{urr}}$]
[$_\varphi$ twe$\overset{*}{\text{e}}$ en d$\overset{\o}{\text{e}}$rtig k$\overset{*}{\text{a}}$sten]

URR: [$_\varphi$ h$\overset{\o}{\text{o}}$nderd$_{\texttt{urr}}$]
[$_\varphi$ v$\overset{*}{\text{i}}$eren vijftig d$\overset{\o}{\text{u}}$izend$_{\texttt{urr}}$]
[$_\varphi$ z$\overset{*}{\text{e}}$ven h$\overset{\o}{\text{o}}$nderd$_{\texttt{urr}}$]
[$_\varphi$ twe$\overset{*}{\text{e}}$ en dertig k$\overset{*}{\text{a}}$sten]

Output: honderd v$\overset{*}{\text{i}}$eren vijftig duizend
z$\overset{*}{\text{e}}$ven honderd twe$\overset{*}{\text{e}}$ en dertig k$\overset{*}{\text{a}}$sten

Due to the length of the expression, an AD-boundary after *duizend* is possible, which would prevent deaccentuation of *duizend* by URR. Conversely, the presence of an accent on both *vieren* and *duizend* would mean that the first conjunct in (41), which forms a coordinated structure, contains two accents, and would therefore trigger the insertion of an AD-boundary by the phrasing rule AD-in-coordination (cf. the end of Section 4.8.1.5). We thus correctly predict the coincidence of an accent on *duizend* and a subsequent AD-boundary. However, as already mentioned, AD-boundaries are not yet known when URR applies. We will further discuss this paradox in Section 7.3.6.4.

7.3.6 Discussion

7.3.6.1 Distribution of focus over words

The text planner indicates non-default focus by attaching the features `foc` (and `back`) to the lemmas in an APS. Since the level of APS is intended to abstract away from syntactic details, function words are normally not specified, because these are assumed to be inserted by SynR. As a result of this abstraction, function words are automatically unfocused and unaccented. For many function words, this prediction is indeed correct. In particular, pure function words, i.e. those that serve a only grammatical function and have no semantic content of their own, cannot be focused and accented.[14] In (42-a), for example, the verb *rekent* subcategorizes for a prepositional object with the preposition *op*. This preposition is purely functional, not contributing any semantic content of its own. In (42-b), we find a passive form of an intransitive verb. Since there is no semantic role available for the syntactic function of subject, its place is filled with the semantically empty element *er*. Example (42-c) shows the verb *vergist*, which obligatory subcategorizes for the reflexive pronoun *zich*. In all three cases, focus on the purely functional element is very unlikely.

(42) a. ?De organisatie rekent $\overset{*}{\text{op}}$ goed weer.
the organisation counts on good weather

b. ?$\overset{*}{\text{Er}}$ wordt gestaakt.
there is striked.
'A strike is going on.'

c. ?De minister heeft $\overset{*}{\text{zich}}$ vergist.
the minister has himself mistaken
'The minister was mistaken.'

The assumption that focus is distributed over the lemmas in an APS goes counter to linguistic theories which assume that focus is distributed over seman-

[14] Barring metalinguistic uses. It is virtually impossible to claim that a certain accent distribution is *ungrammatical*, because for every odd accentuation, one can imagine a context in which it would be appropriate. For example, *g$\overset{*}{\text{e}}$loven* 'belief' has an accent that is not on the main-stressed syllable of the word. Moreover, the accented syllable contains a schwa, which normally cannot be accented. However, this accentuation is appropriate to correct a misunderstanding, as in: *Moet ik dat beloven?* 'Do I have to promise that?'; *Nee, g$\overset{*}{\text{e}}$loven* 'No, believe'.

tic terms (Schmerling 1976; Gussenhoven 1984), and may not be entirely appropriate. At least some function words, like the proximal determiner *die* in (43-b), can be accented, which implies that they can be focused. However, since an APS abstracts away from such function words, there is no appropriate lemma to attach the feature foc to. A partial solution is to specify the function word anyway (43-a).

(43) APS: (bloem<near,back>(DET:<foc>))
Output: die bloem
that flower

This is impossible in (44), where it is not clear whether an additional (V:<foc>) refers to the auxiliary *zou* or to *hebben*. Adding the auxiliary's lemma (i.e. (V:<foc>zullen)) defeats the whole purpose of an APS, that is, abstracting away from purely syntactic matters.

(44) APS: (V:doen<past,fut,perf>[<speak>][<far,top>])
Output: dat zou ik hebben gedaan (ware het niet dat ...)
that would I done have (if only ...)
I would have done that (if only ...)'

In general, the possibility to focus on proximal/distal relations, modality, negation or other aspects appears to require a more semantic form of representation than an APS can offer; see also (Bolinger 1983). We leave this issue for future research.

7.3.6.2 Argument or modifier

As far as deaccentuation of the predicate is concerned, correct application of SAAR crucially depends on the identification of arguments. Clearly, SLG has an advantage over TTS here, as the relevant information is automatically generated, relying on a lexicon that encodes the number and type of arguments of a verb. Still, in a number of situations it is not that obvious how many arguments the lexical entry of a verb should have. If a constituent cannot be omitted, or the verb subcategorizes for its form (e.g. a prepositional object), we can safely assume that it is an argument. Similarly, there are some syntactic tests that may help in diagnosing arguments. However, sometimes we can only rely on semantic intuitions about the 'strong relation' between a verb and a supposed argument. As a consequence, when a predicate is incorrectly accented, we are forced to choose between attributing the error to malfunctioning of SAAR or to an incorrect specification of the argument structure.

Consider example (45-a), where everything except the subject is assumed to be focused. The predicate *gaat* ought to be unaccented. However, because the instrument *met de stoptrein* is supposed to be a modifier, and is focused as well, it prevents deaccentuation of *gaan*. Thus, either this is an exception to SAAR, or the instrument is in fact an argument. Notice, however, that the predicate cannot be accented either with other modifiers (45-b-c), which suggests that some other factor is responsible for the deaccentuation of the predicate. A possible explanation is that *gaan*, which also serves as an auxiliary of future tense in Dutch, is a semantically light verb, as most of the meaning of 'going to' is expressed by the preposition of its argument (*naar* 'to'), rather than by the verb itself. Alternatively, we may conjecture that in the context of train travel descriptions, predicates like *gaan* are contextually activated to such a degree that they are backgrounded.

(45) a.

		*			*		*
(45) a.	?we	gaan	met	de	stoptrein	naar	Amsterdam
	we	go	with	the	fast-train	to	Amsterdam

b.

	*			*		*
?we	gaan	met	de	kinderen	naar	Amsterdam
we	go	with	the	children	to	Amsterdam

c.

	*			*		*
?we	gaan	met	een	slakkegangetje	naar	Amsterdam
we	go	with	a	snail's-pace	to	Amsterdam

Another example concerns the verb *groeien* 'grow', which occurs seven times in the corpus of flower descriptions, but is never accented. In (46), for instance, the subject *de plant* is assumed to be unfocused, and the locative *in venen en natte heide* is taken to be a modifier. According to SAAR, therefore, the predicate must be accented, which is incorrect.

(46)

		*		*		*	*
?de	plant	groeit	in	venen	en	natte	heide
the	plant	grows	in	peats	and	wet	heath

However, it may be argued that the predicate 'grow' in (46) is not used in its literal sense of 'becoming bigger', but rather as a synonym for 'occurring at'. The latter is much more likely to have a locative argument; compare (47).

(47)

			ø		*		*	*
de	plant	komt	voor	in	venen	en	natte	heide
the	plant	occurs	`prt`	in	peats	and	wet	heath

Moreover, the predicate does have an accent in (48), where it can only be interpreted in the sense of 'becoming bigger'.

(48)

		*	*
de	plant	groeit	snel
the	plant	grows	fast

So with respect to (46), we have reason to believe that it is not an exception to SAAR, but an example of how a predicate's argument structure can depend on its interpretation.

7.3.6.3 Accentuation of adverbs

As for any other word, ProsR accents an adverb only if it is focused, which it is either by explicit specification from the text planner or by the default focus rules in SynR. These defaults rely on the content/function word distinction: only content words receive focus. While this distinction is relatively clear for verbs (auxiliaries and copula versus main verbs) and nouns (common and proper nouns versus pronouns, demonstratives, etc.), the situation is not so clear for adverbs. Some adverbs, like *gisteren* ‘yesterday’ and *rechtdoor* ‘straight on’, have a semantic content similar to that of common nouns and main verbs. As a matter of fact, it can be argued for many adverbs that they are adverbially used adjectives. Adverbs of this type can thus be considered content words. Others have an anaphoric function akin to that of pronouns, e.g. *hier* ‘here’, *daar* ‘there’ or *waarna* ‘where-after’, suggesting a classification as function word. For many other adverbs, a classification as either function or content word is difficult. Examples are:

- discourse markers (e.g. *echter* ’however’ and *vervolgens* ‘next’)
- intensifiers (e.g. *zeer* ‘very’ and *nogal* ‘quite’)
- negations (e.g. *niet* ‘niet’ and *nooit* ‘never’)
- particles (e.g. *ook* ‘also’ and *al* ‘already’)

In spite of these difficulties, default focusing of adverbs by SynR is rather simplistic: those that are clearly anaphoric are classified as function words, all others as content words. As a result, adverbs are sometimes incorrectly accented. Some examples from WFD are given in (49): *wel* in (49-a), *ook* in (49-b), and *vaak* in (49-c).[15]

(49) a.

?hij	wordt	wel	anderhalve	meter	hoog
it	becomes	as-much-as	one-and-a-halve	meter	high

b.

?ook	dit	is	een	zeldzame	soort
also	this	is	a	rare	species

c.

?de	plant	heeft	vertakte	#	vaak	liggende	stengels
the	plant	has	branching		often	recumbent	stems

In PTDGen, the default focus is overruled by the text planner, which backgrounds all adverbs. This appears to be the best strategy, as virtually all adverbs in this domain are discourse markers, such as *eerst* ‘first’, *vervolgens* ‘next’, and *tenslotte* ‘finally’. This type of adverbs was found to be rarely accented (van Donzel 1999, p. 154).[16]

Needless to say, there is scope for improvement here. Possibly, a more refined classification of adverbs will help to improve their default accentuation;

[15] The accent on *vaak* causes RR to erroneously delete the accent on *liggende*.

[16] Interestingly, however, discourse markers were often followed by a pause.

see also Hirschberg (1993, p. 321). However, little appears to be known about what actually causes different types of adverbs to be focused/accented (but see e.g. Verhagen (1986) and Gussenhoven (1984, p.125)).

On the one hand, since the problem is one of focus distribution rather than of pitch accent placement, a solution should be expected from discourse theory. In terms of SLG, this amounts to saying that, rather than relying on defaults, the text planner must take responsibility for the focusing of adverbs. On the other hand, there appear to be systematic relations between the interpretation of focus particles (a special class of adverbs) and their accentuation (Verhagen 1986; Elffers 1992; Heinzel 1999). This suggest that the accentuation of adverbs is not 'simply' a matter of correctly predicting their focus, but requires a semantically driven accentuation rule in the spirit of SAAR.

7.3.6.4 Interaction between pitch accent placement and intonational phrasing

We encountered at least two cases that suggest an interaction between pitch accent placement and intonational phrasing. First, we found that an AD-boundary between a predicate and an extraposed argument blocks a predicate's deaccentuation by SAAR (cf. example (15-c)). Second, we found that URR deletes the accent on a unit term followed by another accent, unless an intervening AD-boundary occurs (cf. example (38)). The problem with these cases is that they contradict the assumption that pitch accent placement *precedes* intonational phrasing, on the assumption that correct intonational phrasing partly depends on the pitch accent distribution, as illustrated by the rule AD-in-Coordination in Section 4.8.1.5, where intonational phrasing in coordinated structures was argued to depend on the number of accents in the conjuncts. Other rules use the number of pitch accents as an indication for the weight of an AD. Moreover, the pitch accent distribution must prevent the creation of ill-formed AD's containing no accents, which was found to be one of the flaws in Nespor & Vogel's intonational phrasing algorithm (Section 4.4.2.4). Therefore, reversing the order of pitch accent placement and intonational phrasing is not an attractive option.

The interaction is a rule-order paradox: phrasing requires accentuation, but accentuation requires phrasing. In the case of URR, a possible solution is provided by *precompilation* as proposed by Hayes (1990) and used for the accentuation of numerals in Hungarian in Varga (1998). This implies that, for each unit word, the lexicon generates two forms, one accented and one unaccented, together with their appropriate contexts, and that the choice of the form is delayed until the required context can be checked. Precompilation appears to be intended to handle syntax-dependent phonological effects. Accounting for SAAR by precompilation would force us to assume that the lexicon generates at least two precompiled forms (accented and unaccented) for *every* predicate. An alternative direction to explore is Optimality Theory (Prince and Smolensky 1993), as it abandons the concept of rule ordering altogether (Selkirk 2000). For the time being, the issue is provisionally dealt with in the implementation of ProsRby undoing deletions due to SAAR or URR if an intervening AD-boundary is inserted later on.

7.4 Tune choice

When for each of the generated utterances the locations of pitch accents and intonational boundaries have been determined, the last task left in order to arrive at a complete intonational specification is choosing a tune. This requires two sorts of choices. First, for each pitch accent, we have to choose its type, i.e. H^*L, L^*H, or H^*LH. Second, for each AD, tone morphemes like downstep may be added. Although investigating the whys and the wherefores of tune choice is beyond the scope of the current project, it is clear that some sort of provisional tune choice must be carried out if we are to synthesise our output. In this section, therefore, a simple mechanism for tune choice will be presented (see also the general discussion in Section 9.3.4).

7.4.1 Type of pitch accent

We assume that the default pitch accent is H^*L. Deviation of the default only occurs for the last pitch accent in an AD, which may be either L^*H or H^*LH. The reason for this is that both L^*H and H^*LH end in an H tone, phonetically corresponding to an upwards pitch movement that effectively marks the end of a non-final AD. The choice between either of them is made dependent on the distance of the pitch accent to the following AD boundary. If the accent is on the last word in the AD, L^*H will be used (50-a), but if the accent is on one of the prefinal words in the AD, H^*LH is chosen (50-b).[17] This avoids long, high-pitched stretches of speech, which are hard to synthesise convincingly. There is some empirical evidence that the same strategy is used by human speakers (Terken 1993, p.39).

(50) a.

	L^*H		H^*L	H^*L
de	grote	#	klokvormige	bloemen
the	large		bell-shaped	flowers

b.

	H^*LH						H^*L		H^*L H^*L
de	bloemen	van	de	plant	#	zijn	groot	en	klokvormig
the	flowers	of	the	plant		are	big	and	bell-shaped

If the utterance is a question, as indicated by the feature `polar` or `wh` in its syntactic structure, H^*LH or L^*H is chosen for its last accent, the choice between them being determined by the considerations of length given above. The algorithm for pitch accent selection can therefore be stated in the form of the following three rules.

[17]The intonational boundary in this particular location is for exemplification only, and would actually not be generated by $\mathcal{NEW}$.

(51) **AD-final-accent**

For each final pitch accent in a non-final AD:
if it is on the last word in the AD, select L*H, otherwise, H*LH.

(52) **Question-accent**

For the final pitch accent in a question:
if it is on the last word of the utterance, select L*H, otherwise, H*LH.

(53) **Default-accent**

Select H*L for all other pitch accents.

7.4.2 Boundary tones

Recall that every AD starts with an initial boundary tone. The default, unmarked choice is %L, whereas a %H tone has particular pragmatic implications. For instance, there is experimental evidence that it interacts with the subsequent pitch accent to signal friendliness, politeness, irritation, and aloofness (Grabe, Gussenhoven, Haan, Marsi, and Post 1998). Little is known about the meaning of this *high onset* in Dutch, but it is clear that it is only appropriate in particular situations (Keijsper 1984a; Keijsper 1984b; Pierrehumbert and Hirschberg 1990). In addition, it appears to be related to the discourse structure in the sense that a new discourse segment starts with a relatively high pitch (Wichmann, House, and Rietveld 2000). As far as our application domains are concerned, there is at least one suitable context. Recall that travel descriptions often consist of multiple paragraphs (e.g. from A to B, next from B to C, and finally from C to D). We found that in many cases a %H tone can be used at the start of a new paragraph in order to signal the paragraph break. It appears to highlight the textual structure and increase the liveliness of the synthetic speech.[18] This strategy is formalised in the two rules below.

(54) **Paragraph-marking-initial-boundary-tone**

Select %H for the first AD of a paragraph.

(55) **Default-initial-boundary-tone**

Select %L for all other AD's.

[18] However, if the distance to the first pitch accent is too large, the result is a long, high-pitched stretch of speech, which sounds less natural. Notice that this is similar to our observation regarding the use of an AD-final L*H (cf. Section 7.4.1). In addition, if the first syllable of the AD is accented, selecting a high boundary tone is useless, as there is no space for its phonetic realization, although it does no harm either.

Following Gussenhoven and Rietveld (1992a), we assume that an utterance-final H*L is combined with $\mathrm{L}_{\%}$ and an L*H with $\mathrm{H}_{\%}$. No boundary tone is added if the last pitch accent is H*LH, as its final H tone serves as a boundary tone in AD-final position.

(56) **Utterance-final-boundary-tone**

If the last pitch accent in the utterance is H*L (L*H), then select $\mathrm{L}_{\%}$ ($\mathrm{H}_{\%}$) for the last AD.

7.4.3 Tone morphemes

The only tone morpheme that we will take into account is downstep (see the paragraph on scaling in Section 5.6.2.2). Downstepped contours would appear to be typically used for coordinations. Recall from Section 4.8.1.5 that we hypothesised that there are basically two ways to phrase a coordination: either all conjuncts are taken together in a single AD, or each conjunct is assigned to a separate AD. It was assumed that the choice depends on how the accents are distributed over the coordination: the former is taken for simple coordinations, in which each of the non-final conjuncts carries at most one pitch accent, otherwise the latter is chosen. The same distinction between simple and complex coordinations also plays a role with respect to downstep. It would appear that downstep is appropriate when applied to a simple coordination (57-a), but odd when applied to a complex coordination, at least for non-final conjuncts (57-b). As in intonational phrasing, the final conjunct appears to form an exception, as it can carry multiple accents without blocking downstep (57-c).

(57) a.

![$_{\mathrm{AD}}$	Fr$\overset{*}{\mathrm{a}}$nkrijk	D$\overset{*}{\mathrm{u}}$itsland	en	B$\overset{*}{\mathrm{e}}$lgie]
	France	Germany	and	Belgium

b.

? ![$_{\mathrm{AD}}$	N$\overset{*}{\mathrm{o}}$ord-Fr$\overset{*}{\mathrm{a}}$nkrijk	Z$\overset{*}{\mathrm{u}}$id-D$\overset{*}{\mathrm{u}}$itsland	en	$\overset{*}{\mathrm{O}}$ost-B$\overset{*}{\mathrm{e}}$lgie]
	North-France	South-Germany	and	Eastern-Belgium

c.

![$_{\mathrm{AD}}$	Fr$\overset{*}{\mathrm{a}}$nkrijk	D$\overset{*}{\mathrm{u}}$itsland	en	$\overset{*}{\mathrm{O}}$ost-B$\overset{*}{\mathrm{e}}$lgie]
	France	Germany	and	Eastern-Belgium

In sum, intonational phrasing and downstepping are taken to be complementary options as far as coordinations are concerned. A simple coordination has no internal phrasing, but allows downstep. In contrast, a complex coordination has internal AD-boundaries, but does not tolerate downstep. This is captured in the following rule.

(58) **Downstep**

For each AD containing a simple coordination, apply downstep.

The implementation of (58) requires some extra effort. Phonologically, downstep is a tone morpheme that is attached to an AD-boundary. The start and end of a simple coordination do not necessarily correspond to AD-boundaries. In (59), for instance, the AD starts with *ze*, whereas the actual coordination starts with *groot*. Therefore, the implementation of (58) requires a search for either an AD-boundary coinciding with the start of the coordination, or the first one preceding it.

(59)	![$_{\text{AD}}$ ze	zijn	gr$\overset{*}{\text{o}}$ot	h$\overset{*}{\text{o}}$ekig	en	st$\overset{*}{\text{e}}$vig]
	they	are	big	angular	and	firm

As a matter of fact, some parts belonging to the same AD as the coordination may be accented as well. For instance, (60) has an accent on the verb *uitgestorven*, which is not a part of the preceding coordination, and downstep may be less appropriate here. Alternatively, an extra AD-boundary can be inserted after the coordination, or downstep can be omitted altogether. The former violates the constraint that no boundary should be inserted before the verb cluster; the latter fails to signal the structure of the simple coordination by intonational means. As intuitions are not clear-cut on this matter, we choose to implement the latter alternative in ProsR.

(60)	? ![$_{\text{AD}}$ een	plant	die	in	Fr$\overset{*}{\text{a}}$nkrijk	D$\overset{*}{\text{u}}$itsland	en	N$\overset{*}{\text{e}}$derland	is
	a	plant	that	in	France	Germany	and	Netherlands	is

$\overset{*}{\text{u}}$itgestorven]
died-out
'a plant that has become extinct in France, Germany and the Netherlands'

Although the above approach to downstep in coordinations works satisfactorily, there is another way to render complex coordinations like (57-b), namely, by a special form of downstep that affects only the 'primary' accent in each conjunct. With respect to (57-b), this means that the H* targets for *Noord*, *Zuid*, and *Oost* are downstepped, while those for *Frankrijk*, *Duitsland*, and *Nederland* remain at the same level. This variant can be included into the phonological description as a new tone morpheme, with a corresponding phonetic implementation in the speech synthesiser.

In the same vein, it was found that appositions cannot be optimally accounted for within the current system. These are typically accompanied by a temporary lowering and compression of the register: the overall pitch level drops at the start of the apposition, but returns to the previous level when finished. This could be modelled by a tone morpheme for pitch reduction. As far as the implementation is concerned (cf. the scaling function in Figure 5.16), lowering of the register can be accomplished by decreasing parameter N, and compression of the register by decreasing W.

7.5 Summary

This chapter described ProsR, the module of CONPAS that is responsible for the prosodic realization of utterances on the basis of their syntactic structure, focus distribution, argument structure, and the surface string. The output is a string of words interspersed with prosodic annotations for prosodic constituents (φ's and π's), AD's, pitch accents, boundary tones, and tone morphemes. While producing this output, ProsR carries out the following four tasks:

1. **Derivation of prosodic constituents**
   ProsR implements the new algorithm for φ-phrasing in Dutch proposed in Chapter 3. Apart from the inclusion of the category Num in the definition of adjacent modification, and an extra rule for clitic movement, no substantial modifications were required, as the algorithm produced the intended results. In addition, ProsR implements a simple π-phrasing rule.

2. **Pitch accent placement**
   Starting from the simple algorithm presented in Chapter 2, several improvements were proposed.

   (a) With respect to SAAR, two steps were implemented: initial accentuation according to the focus distribution, followed by deaccentuation of the predicate where appropriate. It was argued that the internal structure of some compounds (phrasal compounds, which may carry two accents, and compound verbs) must be transparent for the accentuation algorithm. Furthermore, a set of focus projection rules was introduced to define the relation between a focused argument and the words it consists of. Finally, it was observed that deaccentuation of the predicate is blocked with topicalised arguments, and with extraposed arguments preceded by an AD-boundary.

   (b) An additional deaccentuation rule was introduced for NP's that serve as NP modifier (NPPD).

   (c) The Lexical Rhythm Rule (LRR) was implemented to account for an small class of exceptional quantifiers and adverbs.

   (d) An investigation of the accentuation of numerals motivated the introduction of the Unit Rhythm Rule (URR). It was shown that the combination of φ-phrasing plus rhythm rules correctly predicts the accent distribution in complex numerals. It was observed that URR is blocked by an intervening AD-boundary.

   The complete algorithm for pitch accent placement is summarised in Table 7.1.

   Some remaining problems were discussed.

   (a) The assumption that focus is distributed over the lemmas in an APS is too restricted, because it does not allow certain cases of narrow focus with an accented function words.

Table 7.1: Rules of the algorithm for pitch accent placement.

Name:	*Definition:*
Sentence Accent Assignment Rule (SAAR)	If focused, every predicate, argument, and modifier must be accented, with the exception of a predicate, that, discounting unfocused constituents, is adjacent to an argument. (Exceptions are possible for topicalised and extraposed arguments; cf. Section 7.3.1.2.)
NP Premodifier Deaccentuation (NPPD)	Delete the pitch accents on an NP that serves as a premodifier of another NP.
Rhythm Rule (RR)	$* \rightarrow \emptyset \;/\; [_{\varphi} \ldots [_{\omega}\ *\ ] \ldots [_{\omega}\ \_\ ] \ldots [_{\omega}\ *\ ] \ldots]$
Lexical Rhythm Rule (LRR)	$* \rightarrow \emptyset \;/\; \left\{ \begin{array}{l} [_{\mathrm{U}} \ldots [_{\omega}\ [_{\omega}\ *\ ]\ [_{\omega}\ \_\ ]\ ]_{\texttt{lrr}} \ldots [_{\omega}\ *\ ] \ldots] \\ [_{\omega}\ [_{\omega}\ \_\ ]\ [_{\omega}\ *\ ]\ ]_{\texttt{lrr}} \end{array} \right\}$
Unit Rhythm Rule (URR)	$* \rightarrow \emptyset \;/\; [_{\mathrm{U}} \ldots [_{\omega}\ \_\ ]_{\texttt{urr}} \ldots [_{\omega}\ *\ ] \ldots]$

(b) The distinction between argument and modifier is not always clear, and appears to be context-dependent for some predicates.

(c) The default accentuation of adverbs on the basis of the content versus function word distinction results in too many accented adverbs. A more refined categorisation of adverbs might be useful. It was additionally suggested that the accentuation of adverbs is not only a matter of correct prediction of their focus by the text planner, but may require a semantically driven accentuation rule.

(d) Examples of interaction between pitch accent placement and intonational phrasing were found with respect to SAAR and URR, leading to a rule-order paradox for which we have no sound solution.

3. **Intonational phrasing**
   ProsR uses the $\mathcal{NEW}$ algorithm for intonational phrasing developed in Chapter 4.

4. **Tune choice**
   A relatively simple algorithm for tune choice was described, which is

Table 7.2: Rules of the algorithm for tune choice.

Name:	*Definition:*
AD-final-accent	For each final pitch accent in a non-final AD: if it is on the last word in the AD, select L*H, otherwise, H*LH.
Question-accent	For the final pitch accent in a question: if it is on the last word, select L*H, otherwise, H*LH.
Default-accent	Select H*L for all other pitch accents.
Paragraph-marking-initial-boundary-tone	Select %H for the first AD of a paragraph.
Default-initial-boundary-tone	Select %L for all other AD's.
Utterance-final-boundary-tone	If the last pitch accent in the utterance is H*L (L*H), then select L% (H%) for the last AD.
Downstep	For each AD containing a simple coordination, apply downstep.

summarised in Table 7.2. It was suggested that a high initial boundary tone can be used to signal the start of a new paragraph, and that downstep is particularly appropriate for simple coordinations. In addition, it was observed that an alternative form of downstep exists, which cannot be generated with the current system, but is in principle not difficult to include in the phonological description and its phonetic implementation. The same goes for the pitch reduction accompanying appositions.