Rhythm versus Analogy: Prosodic Form Variation in Dutch Compounds

Anneke Neijt, Robert Schreuder

Radboud University, Nijmegen, The Netherlands

Key words

Dutch compounds
linking elements
phonological variation
speech production and perception
stress clash

Abstract

Creating compound nouns is the most productive process of Dutch morphology, with an interesting pattern of form variation. For instance, *staat* ‘nation’ simply combines with *kunde* ‘art’ (*staatkunde* ‘political science, statesmanship’), but needs a linking element *s* or *en* to form *staatsschuld* ‘national debt’ and *statenbond* ‘confederation’.

Previous research has shown that the use of linking elements is guided by paradigmatic analogy, a factor that in the absence of other factors would lead to paradigm uniformity. However, there is considerable freedom in the use of linking elements, suggesting that other factors are relevant as well. We present studies showing that both stress and length affect their use, and that, in an experimental setting, the linking element *en* is less favored in lengthened compounds. However, the results observed in this experiment can only be explained satisfactorily in terms of rhythm: the preference for a recurrent pattern of stressed and unstressed syllables.

The general conclusion of this study concerns the distinction between language behavior guided by stored knowledge or by processing factors. Models based on analogy (exemplar-based models) rely on stored knowledge. This study shows that apart from that, rhythm plays its own role. Rhythmic structures facilitate language processing, and a preference for perfect rhythm explains which variant of a compound (with or without the linking element) is chosen. Given the universal nature of analogy and rhythm, the issue of the balance between these two components of linguistic knowledge is relevant for a wide array of languages.

1 Introduction

The function of linking elements

Between the modifier and the head of Dutch nominal compounds a linking element is sometimes used that is homophonous with the plural suffixes *-en* and *-s*. The status

Acknowledgments: We would like to acknowledge the support of Harald Baayen, who urged us to investigate the role of rhythm in the CELEX database and who helped us with the selection of data. Thanks also to the audience of the Morphology Meeting in Vienna, 2004, and at the TIN-dag, 2005, especially Ton van der Wouden. Final thanks go to Andrea Krött, Rik Smits, Judith Hanssen, and the reviewers of *Language and Speech*, for their comments on earlier versions of this paper.

Address for correspondence. Prof. dr. A. Neijt, Radboud University, Dutch Department, PO Box 9103, NL 6500 HD Nijmegen, the Netherlands; email: <a.neijt@let.ru.nl>.
of these linking elements is unclear (Haeseryn, Romijn, Geerts, de Rooij, & van den Toorn, 1997; Krott, 2001). Their form suggests plurality, and in many cases a plural interpretation is indeed most prominent, but many other cases exist in which a singular interpretation seems to be more prominent even though the form of the modifier suggests a plural (See Table 1, upper half). In a similar vein, there are many juxtaposed compounds, that is, forms without linking elements, that receive a singular interpretation in accordance with their form, but again many other cases exist in which a plural interpretation is more prominent even though the form of the modifier suggests a singular (See Table 1, lower half). Such data show that the distribution of linking *en* and *s* cannot be predicted exclusively on the basis of meaning.

**Table 1**

Examples that illustrate matching and mismatching forms and meanings in Dutch compounds

<table>
<thead>
<tr>
<th>plural form</th>
<th>plural meaning</th>
<th>singular meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>huizenrij</td>
<td>‘row of houses’</td>
<td>mannenslem</td>
</tr>
<tr>
<td>huis + en + rij</td>
<td>‘house + en + row’</td>
<td>man + en + stem</td>
</tr>
<tr>
<td>damesorkest</td>
<td>‘women’s orchestra’</td>
<td>varkensoogjes</td>
</tr>
<tr>
<td>dame + s + orkest</td>
<td>‘lady + s + orchestra’</td>
<td>varken + s + oogjes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>singular form</th>
<th>singular meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>schaapherder</td>
<td>‘shepherd’</td>
</tr>
<tr>
<td>schaap + herder</td>
<td>‘sheep + herdsman’</td>
</tr>
<tr>
<td>kleuterklas</td>
<td>‘kindergarten’</td>
</tr>
<tr>
<td>kleuter + klas</td>
<td>‘toddler + class’</td>
</tr>
<tr>
<td>deurmat</td>
<td>‘door + mat’</td>
</tr>
<tr>
<td>deur + mat</td>
<td>‘door + mat’</td>
</tr>
<tr>
<td>kadekant</td>
<td>‘quay-wall’</td>
</tr>
<tr>
<td>kade + kant</td>
<td>‘quay + side’</td>
</tr>
</tbody>
</table>

Even though a strict relation between form and meaning is absent, recent psycholinguistic, computational, and corpuslinguistic investigations found evidence for plurality and some other semantic factors that guide the distribution. Linking *en* in written forms (pronounced as schwa in the standard variant of Dutch) marks plurality of the modifier, as shown in perception experiments (Neijt, Schreuder, & Baayen, 2004; Schreuder, Neijt, van der Weide, & Baayen, 1998). In novel compounds, semantic factors such as animacy and concreteness of the modifier affect the choice of linking elements (Krott, Krebbers, Schreuder, & Baayen, 2002). Plural meaning is a solid factor in production also, as shown in a picture naming task with illustrations that indicate situations of plural or singular meaning (e.g., *bananenballon*, with linking *en*, is used more often for a balloon decorated with drawings of bananas and *banaamballon*, without linking *en*, is used more often for a balloon in the form of a banana; Neijt, Krebbers, & Fikkert, 2002).

Observe that in these Dutch experiments, meaning is investigated, not syntactic behavior, as in the German experiments described by Koester, Gunter, Wagner, and Friederici (2004). They found no traces of plural agreement in auditory processing using event-related brain potentials. The test measured phrases such as *ein*/ *zwei Ohrenzeuge* ‘one/two ear + s witness’ and *ein*/ *zwei Liederabende* ‘one/two song + s nights’, with agreement or disagreement between the determiner and the head of the compound, and, crucially, disagreement or agreement between the determiner and the modifier of the compound. There were no traces of such mismatches in processing,
even though the modifiers in their experiment were homophonous with plural nouns (Ohren and Lieder are plural nouns in German). One of the explanations offered by Koester et al. is that prosodic cues signal the syntactic relation between determiner and modifier in an early stage of recognition and prevent agreement processing. One may expect that linking *en* in Dutch also does not function as an agreement marker in syntax on the basis of the outcome of the German experiments.

There is evidence that morphological structure plays a role in the selection of linking elements. In German and Dutch compounds that consist of three constituents, the position of linking *s* correlates with the structure of the compound, signaling left or right branching structures (Krott, Libben, Jarema, Dressler, Schreuder, & Baayen, 2004). The main factor governing the selection of linking elements in Dutch, however, is claimed to be paradigmatic analogy (Krott, 2001; Krott, Schreuder, & Baayen, 2001, 2002). Novel compounds most likely take the linking element that occurs in existing compounds with the same modifier (either the whole word, its suffix, or its rhyme, in this order of importance), or, to a lesser extent, compounds take the linking element that occurs most often in existing compounds with the same head. For instance, because *huis* ‘house’ is used more often (90 types in the corpus described below, in Study 1, e.g., *huisdeur* ‘front door’) as a modifier in compounds than *huizen* ‘houses’ (seven types, e.g., *huizenbezit* ‘house ownership’), novel compounds most likely will be formed with *huis* and not with *huizen*. And because *rij* ‘row’ attracts most often modifiers with linking *en* (3 types, e.g., *bomenrij* ‘row of trees’, none without a linking element), novel compounds will be formed most likely with this linking element. Paradigmatic analogy thus accounts for the variation found in compounds with modifiers such as *huis* and, to a lesser extent, heads such as *rij*.

The fact that not only the modifier but also the head of the compound influences the selection of linking elements gives rise to a situation that may lead to variation in cases of conflict between the analogy requirements of the modifier and the analogy requirements of the head. Phonological factors are sources of variation as well. It has been shown that there is a preference for the use of linking *en* in stress clash contexts (Neijt, 2003; Neijt, Krebbers, & Fikkert, 2002). Stress clash contexts in compounds are contexts in which main stress of the modifier and main stress of the head are adjacent. This context arises when a compound is formed on the basis of a modifier with final stress (e.g., *kameel* ‘camel’) and a head with initial stress (e.g., *ogen* ‘eyes’; there is no stress clash in the novel compound *kamelenogen* ‘camel eyes’). In such a context the use of linking *en*, which forms an unstressed syllable with a schwa, would create a compound with a more optimal stress pattern.

The point that stress influences the distribution of linking elements has been taken up by Krott (2001, pp.225–226), who accordingly finds significant distributional differences in the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995). She also finds a significant interaction of the position of stress in the modifier and the choice of linking elements, which may imply that the position of the stress in the modifier alone is relevant for the occurrence of the linking element. This is indeed what Krott found in simulation studies. The stress position of the modifier by itself correctly predicts 64% of the compound forms; a low score, however, in comparison to the factors investigated by Krott in earlier studies. Including the stress position of the head in the training set of her simulation studies does not change the prediction accuracy, but
when the model is trained on other aspects of the modifier of the compound, the rime of its final syllable and the suffix of the modifier, the prediction accuracy increases significantly to 91%. Stress is thus overruled by the other factors. Krott’s conclusion is that the factor stress does not reliably affect the occurrence of linking elements, at least not in existing compounds of the kind she selected, for example, compounds containing a modifier with plural forms with the suffix -en.

On the basis of Krott’s study, one may conclude that paradigmatic analogy is a better predictor than stress, and that the effect of stress is restricted to pseudo-compounds or novel compounds such as the ones included in the studies of Neijt, Krebbers, and Fikkert (2002), and Neijt (2003). The issue is important. When paradigmatic analogy and semantics are the only factors explaining the use of linking en, one might conclude that the linking element is a morpheme (as we did in Schreuder, Neijt, van der Weide, & Baayen, 1998). When the phonological factor stress also governs the use of this element, one must conclude that this linking element is both a morpheme and a phoneme. For this reason, linking en has been called tentatively a phonomorpheme in Neijt (2003). The present paper offers evidence for this position. The central issue of this paper, however, does not concern the status of linking elements, but the relation between factors as diverse as rhythm and analogy, and limitations on the role of stored knowledge. Morphological factors might be explained by analogy (storage), but rhythm is an independent structure-assigning component. We will show that rhythm influences the choice of linking elements in the set of existing compounds, and additionally, that rhythm influences the choice of linking elements by language users in contexts where analogy cannot play a role.

Study 1 offers a detailed survey of the distribution of linking elements in nominal compounds with a nominal modifier in the CELEX database. These findings, which are in line with those of Krott (2001), show that the position of stress is a factor determining the use of linking en. Another ordering of the data shows that length, in terms of number of syllables is relevant. In longer compounds, linking s is used more often, and linking en is used less often. We further investigate the factor length in our corpus (Study 2) and in a preference task (Study 3), which reveals that length indeed plays a role. Study 4 shows, however, that it is rhythm, not length, which explains the behavior of the participants in this experiment. The relation between stress and rhythm will be explained below, and is illustrated extensively in Appendix 3. The next two sections of this paper report on the corpus studies, which investigate the role of main stress in the choice of linking elements.

2 Study 1: Main stress in N+N compounds

This study investigates the role of stress in a large set of nominal compounds, as has been performed before by Krott (2001, p. 225). Our analysis differs from Krott’s analysis on two points. First, the analysis is based on detailed information about the number of syllables in heads and modifiers, and second, our set of data is twice as large. We included compounds that occurred only once and compounds with modifiers that allow no plural formation using -en. Krott (2001) excluded such compounds, because the literature claims that linking en is not possible in such cases. However,
85 compounds occur of this type, for example, *komijnekaas* ‘cumin cheese’. Observe that the linking element is written with *e* in these cases, a fact that is irrelevant to this study. We will refer to the linking element with *en* throughout, also when discussing phonological aspects of this linking element, because *en* is its most common spelling, even though standard speakers of Dutch often pronounce only schwa.

2.1 Method

2.1.1 Materials

The CELEX database contains 29,298 different N+N compounds, that is, compounds with a nominal modifier and a nominal head, juxtaposed or separated by linking *e*, *en*, or *s*. This set of N+N compounds includes the CELEX dictionary forms that do not occur in the CELEX text corpora. We used orthographic representations, but we did not distinguish between linking *e* and linking *en*, because this distinction is due to former spelling conventions that need not concern us here. Hereafter, both forms are referred to with *en*.

2.1.2 Procedure

We determined the number of syllables, the position of main stress in both nouns, and of course the linking element (nil, *en*, or *s*). In order to investigate the possible effect of the position of main stress in the selection of linking elements, we defined two classes of compounds. Clash compounds are compounds with stress clash between modifier and head, disregarding the syllable introduced by linking *en*, and neutral compounds are compounds without stress clash in that position, compare the examples under (1), with the position of main stress underlined, bold printing for the linking elements, and literal English translations provided.

(1) a. Compounds with stress clash (*n* = 16177)

*with two syllables disregarding the syllable introduced by* *en*:

*nil*: *huisdeur* ‘house + door’

*en*: *bessenjam* ‘berry + en + jam’

*s*: *broekspijp* ‘trouser + s + leg’

*with four syllables disregarding the syllable introduced by* *en*:

*nil*: *accentteken* ‘accent + sign’

*en*: *bandijtenbende* ‘bandit + en + gang’

*s*: *fatsoensregel* ‘decorum + s + rule’

*with five syllables disregarding the syllable introduced by* *en*:

*nil*: *acetyleneengas* ‘acetylene + gas’

*en*: *evenementenhal* ‘event + en + hall’

etc.
b. Neutral compounds ($n = 13121$)

with three syllables disregarding the syllable introduced by en:

nil: importwijn ‘import + wine’
en: boodschappenjongen ‘messenger + en + boy’
s: jongensboek ‘boy + s + book’

with four syllables disregarding the syllable introduced by en:

nil: atlasvlinder ‘atlas + moth’
en: garnalencroquet ‘shrimp + en + croquette’

with five syllables disregarding the syllable introduced by en:

nil: televisiemast ‘television + mast’
s: identiteitprobleem ‘identity + s + problem’

e tc.

2.2 Results and Discussion

The distribution of linking elements is presented in Figure 1, which shows that juxta-
posed compounds (nil) are in the majority and that linking elements are used more
often in clash contexts (27%) than in neutral contexts (21%). This difference is signifi-
cant, $\chi^2 (1) = 143.1$, $p < .001$. The distribution of en and s in stress clash contexts, 19 and
8% respectively, is the near mirror image of their distribution in neutral compounds,
5 and 15% respectively, $\chi^2 (1) = 1360$, $p < .001$.

Figure 1

The distribution of linking elements (nil, en or s) in clash compounds and neutral
compounds

These data show that linking en occurs more often in clash compounds than in
neutral ones, confirming the hypothesis that linking en is used not only for reasons
of semantics or paradigmatic analogy, as shown by earlier studies (Krott, 2001; Neijt
et al. 2002, 2004; Schreuder et al., 1998), but also in order to avoid stress clashes.

The fact that linking s occurs more often in neutral compounds may be due to
the function of this linking element as a signal of word boundaries. As in English
The phoneme [s] occurs as an “extra” segment at the periphery of Dutch words, compare stro ‘straw’ and arts ‘physician’. These words contain strings of three consonants, whereas onsets and rimes of syllables in the middle of words contain a maximum of only two consonants. Since the extra segment [s] typically occurs at the word’s periphery, an [s] in the middle of compounds may help to determine the boundary between the modifier and the head of a compound (though, admittedly, this signal will not uniquely determine the boundary, [s] being an “extra” segment both at the beginning and at the end of the word). Observe that in neutral compounds, by definition, the position of the boundary between modifier and head cannot be deduced from the distribution of stressed syllables. In this respect, neutral compounds will be more difficult to parse than clash compounds in which adjacent main stresses indicate a boundary between modifier and head. This may be the reason why linking $s$ occurs more often in the set of neutral compounds than in the set of clash compounds.

In order to further illuminate the use of linking elements in clash contexts, we made a detailed overview of clash compounds with modifiers and heads of one to four syllables, see Appendix 1. Sixteen types of clash compounds are distinguished on the basis of number of syllables ($\sigma$), see (2). Main stress is indicated by underlines.

\begin{equation}
\sigma-\sigma \quad \sigma-\sigma \sigma \quad \sigma-\sigma \sigma \sigma \quad \sigma-\sigma \sigma \sigma \sigma \\
\sigma\sigma-\sigma \quad \sigma\sigma-\sigma \sigma \quad \sigma\sigma-\sigma \sigma \sigma \quad \sigma\sigma-\sigma \sigma \sigma \sigma \\
\sigma\sigma\sigma-\sigma \quad \sigma\sigma\sigma-\sigma \sigma \quad \sigma\sigma\sigma-\sigma \sigma \sigma \quad \sigma\sigma\sigma-\sigma \sigma \sigma \sigma \\
\sigma\sigma\sigma\sigma-\sigma \quad \sigma\sigma\sigma\sigma-\sigma \sigma \quad \sigma\sigma\sigma\sigma-\sigma \sigma \sigma \quad \sigma\sigma\sigma\sigma-\sigma \sigma \sigma \sigma
\end{equation}

b. Examples of modifiers and heads

\begin{itemize}
  \item $\sigma$: huis ‘house’
  \item $\sigma\sigma$: bandiet ‘bandit’
  \item $\sigma\sigma\sigma$: avontuur ‘adventure’
  \item $\sigma\sigma\sigma\sigma$: activiteit ‘activity’
  \item $\sigma$: deur ‘door’
  \item $\sigma\sigma$: bende ‘gang’
  \item $\sigma\sigma\sigma$: opera ‘opera’
  \item $\sigma\sigma\sigma\sigma$: misdadiger ‘criminal’
\end{itemize}

The overview in Appendix 1 shows that compounds with heads or modifiers of four or more syllables, the ones printed in smaller case in (2a), are not frequent enough to be included further in the analysis. In the remaining nine types of compounds uneven distributions occur of the numbers and percentages of compounds without linking elements, with $en$, and with $s$. Some general patterns may be formulated for these nine types.

First, absence of a linking element occurs more often in compounds with a monosyllabic modifier (types $\sigma-\sigma$, $\sigma-\sigma\sigma$, and $\sigma-\sigma\sigma\sigma$; 74, 78, and 77% respectively) than in compounds with modifiers of more than one syllable (70% or less). This difference is significant, $\chi^2 (1) = 69, p < .001$. The explanation may be that the use of linking elements is related to their function in signaling word boundaries in compounds. In compounds with a monosyllabic modifier, the stress clash of the first two syllables signals in an early stage of recognition the fact that two words are combined. Since language users tend to rely on salient information at the beginning of words (the Cohort Model,
Marslen-Wilson, 1987) rather than on salient information in the middle or at the end of words, this effect is found in compounds with monosyllabic modifiers and not, for instance, in compounds with monosyllabic heads.

As a second pattern, in compounds with mono- or disyllabic heads and modifiers, *en* is equally frequent or more frequent than *s*, whereas in compounds with words of more syllables, *s* is more frequent than *en*. This difference between short compounds and long ones is significant, $\chi^2 (1) = 74.5, p < .001$. This length pattern observed for clash compounds in the data of Appendix 1 suggests that the distribution of linking elements is influenced by the length of the compound, counted in number of syllables. Study 2 investigates this effect in the larger set of clash and neutral compounds.

### 3 Study 2: Length in N+N compounds

Observe that use of the linking *en* adds a syllable to the compound, whereas use of the linking *s* does not affect length. In this section, we investigate the possibility that the distribution of linking elements is not only governed by semantics, and paradigmatic analogy, as shown by previous studies, and stress, as shown in Study 1, but also by the length of the compounds. When length is a factor determining the form of compounds, we expect to find *en* more often in shorter compounds than in longer compounds.

#### 3.1 Method

**3.1.1 Materials and procedure**

The 29,298 N+N compounds are described in Study 1. We grouped these compounds according to their number of syllables, disregarding the syllable introduced by linking *en*.

#### 3.2 Results

The distribution of linking elements is presented in (3) and Figure 2.

(3) All compounds, the distribution of linking elements (%), and examples

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>two syllables (n = 10183)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>nil</td>
<td>aal + bes</td>
<td>‘eel + berry, currant’</td>
</tr>
<tr>
<td></td>
<td>en</td>
<td>band + en + pech</td>
<td>‘tire + en + breakdown, tire trouble’</td>
</tr>
<tr>
<td></td>
<td>s</td>
<td>tijd + s + beeld</td>
<td>‘time + s + picture, portrait of an era’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>three syllables (n = 10316)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>nil</td>
<td>tijd + verdrijf</td>
<td>‘time + diversion, pastime’</td>
</tr>
<tr>
<td></td>
<td>en</td>
<td>aardbei + en + tijd</td>
<td>‘strawberry + en + time, strawberry time’</td>
</tr>
<tr>
<td></td>
<td>s</td>
<td>aandacht + s + veld</td>
<td>‘attention + s + field, area for attention’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>four syllables (n = 5297)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>nil</td>
<td>taal + register</td>
<td>‘language + register, register’</td>
</tr>
<tr>
<td></td>
<td>en</td>
<td>abrikoos + en + boom</td>
<td>‘apricot + en + tree, apricot tree’</td>
</tr>
<tr>
<td></td>
<td>s</td>
<td>liefde + s + appel</td>
<td>‘love + s + apple, love apple’</td>
</tr>
</tbody>
</table>

*Language and Speech*
five or more (n = 3502)

- 75% nil aangifte + formulier ‘notification + form, tax form’
- 5% en artisjok + en + bodem ‘artichoke + en + bottom, artichoke heart’
- 20% s bedrijf + s + analyse ‘company + s + analysis, company analysis’

Figure 2
The distribution of linking elements (nil, en or s) in compounds of two, three, four, and five or more syllables

3.3 Discussion
The data in (3) and Figure 2 show that the distribution of nil compounds is not as variable (74, 79, 76, and 75%) as the distribution of compounds with a linking element. As expected, the corpus contains more linking en in shorter compounds. Moreover, the distribution of en and s shows opposite trends. En is used more often in shorter compounds than in longer compounds (short vs. long: 22, 10, 8, and 5%), and s is used more often in longer compounds than in shorter compounds (long vs. short: 20, 17, 11, and 5%). The decrease of en in longer compounds is significant (test for trends in proportions $\chi^2(1) = 887, p < .001$) as is the increase of s ($\chi^2(1) = 815, p < .001$).

The opposite trends in the distribution of en and s can be explained in mutually not exclusive ways. First, shorter words in Dutch tend to be native words, and longer words tend to be non-native words. It may be that morphology and phonology of these two strata differ such that linking en occurs more often in native words and linking s occurs more often in non-native words. Another explanation could be related to the morphological structure of short and long words. Observe that the internal morphological structure of heads and modifiers is variable: some are monomorphemic nouns (tijd ‘time’, abrikoos ‘apricot’), others are nouns derived by a prefix (ver- in verdrijf ‘diversion’) or by a suffix (-de in liefde ‘love’), and still others are frozen compounds which are considered noncompounds by CELEX (an example is aardbei ‘strawberry’, from aarde ‘earth’ and the old form bei ‘berry’). Short words will often be morphologically simplex words, whereas long words are more often derived words. Moreover, it is well known that several suffixes exist that co-occur with linking s. It could well be that such morphological distinctions explain the opposite trends in the distribution of en and s.
Figure 2 suggests the existence of a general tendency to avoid particularly long or short items, given the possibility of the choice in their formation. We leave the explanation of the increase of $s$ found in this figure for future research. We expect that the explanation for the increase of linking $s$ is not related to the explanation for the decrease of linking $en$, given that both kinds of linking elements are phonologically quite different. In the remainder of this paper, we return to study $en$, now from the point of view that language users may prefer short words above longer ones in terms of number of syllables.

4 Study 3: Length experiment

The fact that linking $en$ adds a syllable to the compound might explain why this linking element occurs less often in longer compounds. Language users might simply prefer short words above long ones, as illustrated by clippings such as info and uni for informatie ‘information’ and universiteit ‘university’, which are quite popular in Dutch. We investigate this hypothesis of a relation between the number of syllables and the choice of linking $en$ in an experimental setting, using compounds which vary only in length. We ruled out the possibility that other factors influence the outcome of this experiment by constructing compound pairs with exactly the same modifier, and by the selection of compounds in which $en$ carries no additional meaning that is not already provided by the modifier and in which $en$ is used in free variation.

4.1 Method

4.1.1 Materials

We selected 30 existing compound pairs such as grondstof+leverancier — grondstof+$en$+leverancier ‘material($en$) provider’, ‘provider of raw materials’ that are used in free variation with or without $en$. Pairs with meaning oppositions such as cel+$structuur$ ‘structure within a cell’ and cel+$en$+$structuur$ ‘structure created by cells’ were not included. In order to establish that both variants are actually used, we counted the number of results when the word was entered into Google (cf. Appendix 2). The Google data form an underpinning of our intuition that the compounds included in the test are variable compounds. We call these compounds the basic compounds, because they form the point of departure for the construction of longer compounds by recursive use of compounding. For instance: the basic compound in which the linking element is embedded grondstof($en$)leverancier ‘providers of raw materials’ is used as modifier in the lengthened compound grondstof($en$)leveranciers+$bijeenkomst$ ‘conference of providers of raw materials’. This way, the morphological context of the linking element is similar in basic or lengthened compound. Longer compounds are created by mere addition of a head. The full list of experimental data, including glosses and number of occurrences in Google is presented in Appendix 2.

The internal morphological structure of the basic compounds is variable, that is, we used not only N+N compounds in this experiment (as we did in Study 1 and 2), but also compounds with a more complex morphological structure. For instance
the basic compound *pannenkoekenmix* ‘pan+en+cake+en+mix, pancake mix’ is a left branching ((N+N)+N) compound. The N+N compound *pannenkoek* forms the modifier of the compound *pannenkoek(en)mix*, in which the linking element is optional (it is not optional in *pannenkoek*). Similarly, variation occurs in the internal morphological structure of the heads with which these basic compounds are transformed into lengthened ones. Sometimes a noncompound head is added (for instance: *verpakking* ‘wrapping’) and sometimes a head which itself is a N+N compound (for instance: *inktdoos* ‘ink pad’). Even though the morphological structure of the heads varied, the number of syllables added to the basic compounds is relatively stable: most often two or three syllables are added to the basic compounds, in one instance one syllable, and in two instances four syllables.

4.1.2 Participants and procedure

Thirty-six first year students of Dutch at Radboud University Nijmegen took part in the experiment. Either a basic compound pair or a lengthened compound pair was presented to the participants, who were asked to express their preference for the variant with or without *en* on a seven-point scale. Because fast readers may overlook the presence of *en*, all compounds without *en* were presented at the left hand side of the page, and those with linking *en* were presented at the right hand side. See (4).

(4) Example of data presentation

*basic compound pairs*

| grondstofleverancier | 1 2 3 4 5 6 7 | grondstoffenleverancier |

*lengthened compound pairs*

| grondstoffenleveranciersbijeenkomst | 1 2 3 4 5 6 7 | grondstoffenleveranciersbijeenkomst |

The instruction was to circle number 1 in case the participant was certain to always use the form at the left-hand side, and to circle number 7 in case the participant was certain to use the form at the right-hand side. Number 4 should be circled in case the participant thought that he or she used both variants equally often, and the participants were instructed to circle the other numbers in case they had a preference in one of both directions.

The students rated two lists of 30 items, 15 lengthened and 15 basic compound pairs, in two subsequent weeks. Those who rated the lengthened version in the first week saw the basic version in the next week, and those who rated the basic version in the first week rated the lengthened version in the next week. The compounds were presented in semirandom order, in eight different lists. The students did not receive payment for their participation. The task took about eight minutes per list of 30 items.

4.2 Results and Discussion

As expected, the participants assign higher ratings in basic compound pairs (mean 3.7) than in lengthened compound pairs (mean 3.5). This difference reaches significance in the analysis by participants, \( t(35) = 2.15, p < .02 \); 1-tailed analysis, as well as in the
analysis by items, $t(29) = 1.79, p < .04$. Paradigmatic analogy of the left constituent cannot explain this result, since the same left constituent is used in both conditions. Paradigmatic analogy of the constituent to the right of the linking element cannot explain this result either, since the noun adjacent to the linking element is the same. This noun and the noun used to lengthen the compound, for example, `leverancier+s bijeenkomst` in `grondstof+en+leverancier+s+bijeenkomst`, leads to a combination that will not be used often enough as a head in lexicalized compounds to be stored in the mental lexicon of speakers of Dutch. Therefore, the notion of paradigmatic analogy cannot be used to explain the results.

Although inspection of the data shows that the expected preference of using `en` in basic, not lengthened, compounds is present in 20 compounds included in the test, there are 10 compounds (see Appendix 2) that show an opposite effect. Closer inspection of the full data set shows that both the number of syllables added (most often 2 or 3), and relative length, that is, the number of syllables of the head used to lengthen the basic compound divided by the number of syllables of the basic compound, show no significant correlation with `en` preferences ($p > .25$). A more adequate explanation of the data is developed in the next study.

**5 Study 4: Rhythm**

The data of Study 2 and 3 suggest a possible relation between length of the compound and preference to use linking `en`. The data of Study 3, however, do not reveal a significant item-by-item correlation between the number of syllables and `en` preferences. Thus, it might be that a factor other than length is relevant. In search of an explanation, we made prosodic analyses of the test items in terms of rhythm. Rhythm is a patterned recurrence of events in time, used in speech to form chunks of information that are easy to produce and perceive (Nooteboom, 1997, p. 668; Zec, 2006; and see Riad, 2006, for a general introduction). Since rhythmic patterns based on stress intervals are found in a diversity of languages, Dauer (1983) concludes that rhythm might be a language universal property.

It is not unlikely that the participants, even in this pencil-and-paper test, unconsciously created rhythmic patterns in order to grasp the meaning of the compounds or in order to compare the forms without linking `en` to the forms with linking `en`. Notice that linking `en` is pronounced with a schwa that creates a reduced syllable, which might trigger the relation with rhythm. It has been shown that schwa-insertion in simplex words follows rhythmic patterns in Dutch, see Kuijpers and van Donselaar (1997).

Rhythm creates constituents of two or three elements in, for instance, the digit pronunciation of telephone numbers. The number 3123612448 will be split into 312 - 36 - 124 - 48, or 31 - 23 - 61 - 24 - 48, or any other pattern of two or three syllables. Such patterns are realized by prominence on the final element of a group, as in phonological phrases, compare (5), in which macrons (⎯) indicate more prominence than breves (∪).

(5) Rhythmic structure of strings of numbers

| 312 - 36 - 124 - 48 | ∪∪/∪∪/∪/∪ |
| 31 - 23 - 61 - 24 - 48 | ∪/∪∪/∪/∪/∪ |
The patterns in (5) can be realized by tapping on the syllables indicated by macrons. These syllables form beat positions and the syllables indicated by breves are lapses in between beats. A beat pattern is optimal when it obeys the conditions in (6):

(6) Perfect Rhythm
   a. Recurrence: a pattern contains at least two beats
   b. No-Clash: beat positions are not adjacent
   c. Lapses: no more than two lapses occur between the beats

We assume that (6a) and (6b) are more salient aspects of rhythm than (6c), because lapses between beats can be eliminated by syllable deletion or strengthening of a syllable in a string of more than two lapses (Nespor & Vogel, 1989). The string —∪∪∪— can be realized by deletion as —∪∪— and by strengthening as —∪—∪—.

5.1 Method
The test items were analyzed in terms of bracketed metrical grids (cf. Cinque, 1993; Hayes, 1995; Nespor & Vogel, 1986; Selkirk, 1984, and for Dutch, Neijt, 1999). The brackets reflect the prosodic hierarchy of smaller domains (syllables and feet) embedded in larger domains (prosodic words and compound words). The asterisks indicate the syllable with main stress in a given domain. We illustrate the construction of such representations with bloem(en)bakontwerp ‘flower box design’, one of the compounds included in the test.

(7) compound word (* *)
    compound word (* *)  (*) (*)
    pros. word  (*) (*) (*) (*)
    foot (*) (*) (*) (*) (*)
    syllable (*) (*) (*) (*) (*)
    bloem en bak ont werp

Structure (7) is generated as follows. The syllables form the input of the structure, as indicated by the asterisks assigned to each syllable in the first layer above the words. At the second layer, syllables are grouped into feet. The two syllables of bloemen form one foot, the syllable of bak forms a foot on its own, and the two syllables of ontwerp form two feet. Foot construction is weight sensitive, that is, heavy syllables such as ont and werp form a monosyllabic foot, but a light syllable such as the linking en forms a foot together with the preceding syllable. Feet are head-initial, for example, within a foot, the first syllable gets stress and the other syllable is unstressed.

Feet form head-final prosodic words. This rule applies to the two feet of ontwerp, and vacuously in case the word consists of only one foot, as in bloemen, or in monosyllabic words such as bak. The words bloemen and bak form a compound word with main stress on the modifier. Observe that at this layer, the metrical structure of ontwerp is repeated, such that it is present as input of the top layer in which ontwerp forms
a compound word together with bloemenbak. This optional step in the derivation is called Beat Addition by Hayes (1995, p. 373). For the sake of consistency, we applied Beat Addition throughout the data set. An additional reason for the application of this rule is the fact that the test situation implies a careful word-by-word pronunciation, with secondary stresses fully realized and without reduction of syllables. Hence, a representation with as many grid marks as possible is adequate in this situation.

At the top layer, the compound bloemenbak and the word ontwerp form a nominal compound, with main stress on the modifier. Structure (7) thus describes the fact that main stress in bloemenbakontwerp is on the first syllable, as well as the fact that main stress in bloemen is on the first syllable, and in ontwerp it is on the final syllable.

Rhythmic patterns are based on peaks and valleys, as suggested in Hayes (1995). A syllable dominated by a column higher than the columns of adjacent syllables forms a beat, whereas the syllables in between these peaks of stress form the lapses. It so happens that this algorithm generates peaks and valleys only in our set of test items, no instances with two adjacent highest columns between valleys. All structures thus obey Condition (6b) No-Clash. Without the use of context sensitive prosodic adaptation rules, such as the Rhythm Rule (Kager & Visch, 1988) we thus generated one rhythmic pattern for the test items. In (7), this is the rhythmic pattern with the three beats for the three words: blöemēnbākōntwērp (the rhythmic pattern is indicated by macrons and breves).

Next, consider the prosodic structures of the three variants of bloemenbakontwerp used in the experiment. Observe that the position of macrons and breves in these examples correlates with peaks of columns of beats.

(8) a. blöembākōntwērp

| compound word | (* |   ) |
| compound word | (* | * |   | * |
| pros. word    | (* | * |   | * |
| foot          | (* | * | * |   | * |
| syllable      | (* | * | * |   | * |

blöem  bāk  ōnt  wērp

---

1 The rhythmic interpretation of metrical grid structures is complex. According to Lerdahl and Jackendoff (1983), each layer of the grid represents a potential rhythmic pattern and the metrical grid forms a set of rhythmic realizations, not a single one. For instance: The metrical grid in (7) can be tapped with five beats for each of the five syllables, with four beats for the four feet, with three beats for the three words (— ∪ — ∪ —), with two beats for the compound and the final word (— ∪ ∪ ∪ ∪) or with only one beat, on the first syllable. For this study, it is impracticable to evaluate sets of rhythms. Furthermore, we doubt the validity of the hypothesis by Lerdahl and Jackendoff that in language, like in music, a full set of rhythms is relevant. Rather, we think that language users select one rhythmic interpretation per context or situation.
The forms (8a) and (8b) are rhythmically different from (7), but equally well-formed. Example (8c), however, is too short. It violates Recurrence (6a), which is indicated by the hash sign.

Comparison of the basic compounds (8b) and (8c) with their lengthened variants (8a) and (7) reveals an asymmetry. In the basic pair, the form #blōembāk is ill-formed because it violates Condition (6a), but the alternative form with the linking element blōemēnbāk is well-formed. Both variants of the lengthened compound pair, however, are well-formed (cf. blōemēnbākōntwērp and blōembākōntwērp). Along the lines sketched in (7) and (8), we made grid structures of the test items of Study 3.

### 5.2 Results

The prosodic analysis of the test items (presented in full detail in Appendix 3) shows that two groups of items can be distinguished on the basis of (6a). In a group of nine test items, the basic compounds improve by insertion of the linking element en, compare (9). For instance, blōemēnbāk is in accordance with all three conditions of Perfect Rhythm, but #blōembāk violates Recurrence (6a) (and its logical alternative #blōembāk, not generated by the algorithm we used, violates No-Clash (6b)).

(9) Basic compounds which improve by insertion of linking en

<table>
<thead>
<tr>
<th>not rhythmic</th>
<th>rhythmic</th>
</tr>
</thead>
<tbody>
<tr>
<td>— ∪</td>
<td>— ∪ —</td>
</tr>
<tr>
<td>#blōem + bāk</td>
<td>blōem + ēn + bāk</td>
</tr>
<tr>
<td>#brīev + hōofd</td>
<td>brīev + ēn + hōofd</td>
</tr>
<tr>
<td>#vīs + vōer</td>
<td>vīs + ēn + vōer</td>
</tr>
<tr>
<td>— ∪ ∪</td>
<td>— ∪ —</td>
</tr>
<tr>
<td>#bōek + bīndērs</td>
<td>bōek + ēn + bīndērs</td>
</tr>
</tbody>
</table>
The lengthened counterparts of (9) are not improved by the insertion of the linking \textit{en}, compare (10) below. Forms with violations of Lapses (6c) are indicated by tildes, and two tildes indicate longer stretches of lapses. We assume that these are irrelevant for the reasons discussed above (e.g., lapses in a rhythmic string can be repaired by deletion or strengthening of syllables).

(10) Lengthened counterparts of (9)

\begin{verbatim}
blōem + bāk + ŏntwērp
brīev + hōofd + vōorschrīft
vīs + vōer + vērkōpēr
bōek + bīndērs + ōplēidīng
kīler + hāngēr + vērzāmēling
pān + dēksēl + hāndvāt
dōrn + grōoothāndēl + vērgūnṇīng
gōrdījn + mārk + stāndplāat
kāntōor + āanbōd + tōēnāmē
\end{verbatim}

In the other 21 compound pairs, no violations of Type (6a) occur; neither in the basic variants, nor in the lengthened variants. The 10 forms running counter to the hypothesis are all included in this set of 21 items.

Statistical analyses show a significant effect of the experimental manipulation in the set of nine test words of (9) and (10), but not in the other compounds. Participants use higher rates for \textit{en} in the basic compounds of Set (9) than in their lengthened compound variants (10) (mean 3.4 vs. mean 2.95). The difference is significant (analysis by participants: \(t(35) = 4.03, p < .001\); analysis by items \(t(8) = 3.84, p < .003, 1\text{-tailed}\). In the remaining set of 21 forms without rhythmic violations, no difference between basic and lengthened compounds is found (resp. mean 3.81 and 3.78; analysis by participants \(t(35) = 0.4, p = .35\); analysis by items \(t(20) = 0.31, p = .38\).
5.3 Discussion

As becomes clear from the prosodic analysis of the test items in Appendix 3, a rich array of morphological structures was present in our experimental data. One might therefore legitimately raise the question whether morphological differences between the two sets of items can be held responsible for the results. This possibility can be ruled out on the following grounds.

First, the context in which the linking element is inserted is similar in basic and lengthened compounds. For instance: *grondstof(en)leverancier* is compared with *grondstof(en)leveranciersbijeenkomst*. The linking element in both compounds is inserted between *grondstof* and *leverancier*. Secondly, similar structures occur in the set of compounds that show an effect of rhythm (Set (i) in Appendix 3) and the set of compounds that do not show this effect (set (ii) in Appendix 3). For instance, the following structures occur in both sets (here, the 3 dots … indicate words that form compounds):

(11) Left-branching ( ( … en … )...) compounds:

   (ia) ((boek + en + binders)(op + leiding)) ‘bookbinder training’

   (ib) ((dier + en + verzorgers)(hand + boek)) ‘animal-attendant handbook’

   (iir) ((woning + en + bestand) omputer) ‘housing-stock computer’

Pairwise grouped (( … en …)( …+...)) compounds:

   (iir) ((woning + en + bestand) omputer) ‘housing-stock computer’

   (iir) ((woning + en + bestand) omputer) ‘housing-stock computer’

Center embedded compounds of five elements ((( … en ( …+...))( …+...)):

   (ia) ((boek + en + binders)(op + leiding)) ‘bookbinder training’

   (ib) ((dier + en + verzorgers)(hand + boek)) ‘animal-attendant handbook’

   (iir) ((woning + en + bestand) omputer) ‘housing-stock computer’

These examples illustrate that differences or similarities in morphological structure do not influence the use of linking *en*. Rather, clashes of stressed syllables distinguish contexts in which *en* is preferred from contexts in which the choice is neutral, compare the underlined syllables of the relevant words in (11).

The relation between length and rhythm, and the explanation proposed in terms of rhythm may be further clarified in the following terms. The experiment presented the pairs of compounds in isolation. For Perfect Rhythm a beat pattern is required, that is, at least two beats. The basic compounds in (8) obtain this pattern by insertion of *en*, for example, *bloembāk > bloemēnbaŋk* (underlines indicate the position of word stress; macrons and breves above the vowel letters indicate rhythm, the position of peaks and valleys in grid structures). The lengthened compounds in (10) do not require an extra syllable. The clash of main stress is also present in the lengthened forms, *bloembākōntwērpg*, but the syllable with main stress is simply interpreted as a weak node in the rhythmic string, which is possible, because the word added to the compound provides the following beat for a string that meets the Recurrence requirement of Perfect Rhythm.
Contrastively, the basic pairs of the remaining 21 items reach perfect rhythm without the use of *en*. Observe that the addition of *en* in these cases may create a longer stretch of lapses. Most often, this results in an equally perfect pattern (both —∪— and —∪∪— are well-formed), or a similar string of lapses in the basic and the lengthened compound pair.

The relation between length in terms of syllables and rhythm can now be explained as follows. Perfect Rhythm not only implies strings of syllables without clashes, but also a minimal number of beats. Rhythm is based on recurrence, and therefore, at least two beat positions need to be present. In this way, length plays a role in the choice of linking elements, though as part and parcel of the crucial factor rhythm. Since the test items were presented in isolation, the results of these experiments may be different when the compounds are embedded in larger contexts on the basis of which other rhythmic patterns may be realized. We leave this issue for further study.

Thus, our experiment showed that length is not an independent factor influencing the choice of linking elements. Rather, the results should be explained in terms of rhythm, a factor that has been shown by previous research of natural languages to be relevant for schwa insertion and deletion (Kuijpers, & van Donselaar, 1997). Moreover, the present study shows the relevance of rhythm in the use of linking *en* in Dutch compounds next to analogy, because influences of paradigmatic analogy in our experimental setting have been ruled out.

### Conclusion

The study of linking elements in Dutch compounds in general, illuminates the relation between aspects of language as diverse as plurality (and other semantic aspects), morphology, analogy, and length. The present study contributes to our understanding of the role of the factors stress, length, and rhythm. Our first study, using corpus data, shows that stress patterns, more specifically stress clashes, influence the distribution of linking elements in Dutch N+N compounds. Linking elements occur more often in clashing compounds than in neutral compounds. Moreover, if linking occurs, *en* is more frequent in clashing compounds than *s*, and conversely, *s* is more frequent in nonclashing compounds than *en*.

Our second study shows length to be a factor too. In the corpus data, the two linking elements of Dutch show opposite trends as a function of length measured in number of syllables.

These results might be taken to mean that presence of a linking element in a novel compound depends on paradigmatic analogy, on the basis of stored forms in the mental lexicon with similar lengths or stress patterns. We therefore asked participants in a third study to rate their preference for a linking element in compounds that show free variation, that is, compounds that have been found both with, and without a linking element in text corpora. We also created novel, longer compounds from these compounds by adding another right-hand constituent. The original compound, now the leftmost constituent of the new, longer form, remained exactly the same, so that the role and relevance of paradigmatic analogy remained unchanged.
It turned out that participants were less likely to prefer the presence of a linking element in the left constituent of the lengthened compounds than they had been with respect to the original compounds. This shift cannot be explained by analogy, since both the original and the lengthened compounds were identical. A reanalysis of the data of the third study, furthermore, showed preferences for linking elements to be governed not by length as such, but rather by a prosodic property, rhythm. Irrespective of length, if adding a linking element leads to perfect rhythm (minimally two beats, no clashing beat positions, never more than two contiguous lapses), a linking element will be preferred.

In actual language behavior, the weight of different factors such as analogy, length, stress, and rhythm may vary according to context, style, and genre. Paradigmatic analogy is the important operationalization of all the stored knowledge about forms and phrases gathered by language users up to the moment of a particular instance of language production or perception. It predicts much of the variation we find, but by no means all. In the absence of other factors, analogy builds ever strengthening preferences for ever more frequent forms (structure as emergent in usage-based models, cf. Bybee, 2001), and paradigmatic uniformity would inevitably ensue. In reality, however, variation remains, indicating that independent contextual factors are involved.

The use of linking elements in Dutch is an example of the interplay of stored knowledge, intended meaning, and linguistic context. Apart from analogy, they serve to disambiguate the morphological make-up of especially short compounds, and most importantly to provide compounds with an optimal rhythm, thus presumably facilitating production and perception of speech. We conclude that form variation in Dutch compounds is largely governed by two universal aspects of language: paradigmatic analogy and rhythm. These may explain similar variation phenomena in other languages.

References


 Foris Publications.
KOESTER, D., GUNTER, T. C., WAGNER, S., & FRIEDERICI, A. D. (2004). Morphosyntax,
prosody, and linking elements: The auditory processing of German nominal compounds. Journal of Cognitive Neuroscience, 16, 1647–1668.
on linkers in Dutch noun-noun compounds. Folia Linguistica, 36, 7–22.
KROTT, A., LIBBEN, G., JAREMA, G., DRESSLER, W., SCHREUDER, R., & BAAYEN, H.
(2004). Probability in the grammar of German and Dutch: Interfixation in triconstituent
compounds. Language and Speech, 47, 83–106.
the choice of linking morphemes in Dutch. Linguistics, 39, 51–93.
KROTT, A., SCHREUDER, R., & BAAYEN, R. H. (2002). Analogical hierarchy: Exemplar-
based modeling of linkers in Dutch noun-noun compounds. In R. Skousen, D. Lonsdale,
& D. B. Parkinson (Eds.), Analogical modeling: An exemplar-based approach to language
epenthesis and schwa deletion in Dutch. Language and Speech, 41, 87–108.
MA: The MIT Press.
H. Frauenfelder & L. Komisarjevsky Tyler (Eds.), Spoken word recognition (pp.71–102).
the Netherlands, 1999 (pp.149–161). Amsterdam/Philadelphia: John Benjamins Publishing
Company.
H. van Riemsdijk (Eds.), Germania et alia. A linguistic webschrift for Hans den Besten.
<http://odur.let.rug.nl/~koster/DenBesten/contents.htm>
NEIJT, A. (2005). Abracadabra, the relation between stress and rhythm. In H. Broekhuis, N. Corver,
R. Huybregts, U. Kleinhenz, & J. Koster (Eds.), Organizing grammar. Linguistic studies in
NEIJT, A., KREBBERS, L., & FIKKERT, P. (2002). Rhythm and semantics in the selection of
linking elements. In H. Broekhuis & P. Fikkert (Eds.), Linguistics in the Netherlands, 2002
on interpretation. In L. Cornips, & J. Doetjes (Eds.), Linguistics in the Netherlands, 2004
Bok-Bennema & R. van Hout (Eds.), Linguistics in the Netherlands, 1992 (pp.185–196).
Appendix 1

Distribution of linking elements in clash compounds (Study 1). The number and the percentages of nil, en and s after modifiers of one to four syllables followed by a head of one to four syllables. S and W indicate syllables with (S) or without main stress (W). Percentages for totals < 100 between parentheses.

<table>
<thead>
<tr>
<th></th>
<th>S−S</th>
<th>S−SW</th>
<th>S−SWW</th>
<th>S−SWWW</th>
</tr>
</thead>
<tbody>
<tr>
<td>nil</td>
<td>7496</td>
<td>74%</td>
<td>2322</td>
<td>78%</td>
</tr>
<tr>
<td>en</td>
<td>2187</td>
<td>21%</td>
<td>480</td>
<td>16%</td>
</tr>
<tr>
<td>s</td>
<td>494</td>
<td>5%</td>
<td>177</td>
<td>6%</td>
</tr>
<tr>
<td>total</td>
<td>10177</td>
<td>2979</td>
<td>191</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>WS−S</th>
<th>WS−SW</th>
<th>WS−SWW</th>
<th>WS−SWWW</th>
</tr>
</thead>
<tbody>
<tr>
<td>nil</td>
<td>1068</td>
<td>70%</td>
<td>403</td>
<td>66%</td>
</tr>
<tr>
<td>en</td>
<td>221</td>
<td>15%</td>
<td>64</td>
<td>11%</td>
</tr>
<tr>
<td>s</td>
<td>234</td>
<td>15%</td>
<td>142</td>
<td>23%</td>
</tr>
<tr>
<td>total</td>
<td>1523</td>
<td>609</td>
<td>81</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>WWS−S</th>
<th>WWS−SW</th>
<th>WWS−SWW</th>
<th>WWS−SWWW</th>
</tr>
</thead>
<tbody>
<tr>
<td>nil</td>
<td>205</td>
<td>60%</td>
<td>101</td>
<td>63%</td>
</tr>
<tr>
<td>en</td>
<td>67</td>
<td>19%</td>
<td>31</td>
<td>19%</td>
</tr>
<tr>
<td>s</td>
<td>72</td>
<td>21%</td>
<td>29</td>
<td>18%</td>
</tr>
<tr>
<td>total</td>
<td>344</td>
<td>161</td>
<td>14</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>WWWS−S</th>
<th>WWWS−SW</th>
<th>WWWS−SWW</th>
<th>WWWS−SWWW</th>
</tr>
</thead>
<tbody>
<tr>
<td>nil</td>
<td>23</td>
<td>(62%)</td>
<td>13</td>
<td>(59%)</td>
</tr>
<tr>
<td>en</td>
<td>4</td>
<td>(11%)</td>
<td>5</td>
<td>(23%)</td>
</tr>
<tr>
<td>s</td>
<td>10</td>
<td>(27%)</td>
<td>4</td>
<td>(18%)</td>
</tr>
<tr>
<td>total</td>
<td>37</td>
<td></td>
<td>22</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 2

The compounds included in the experiment (Study 3). Basic compounds: the variable linking element *en* is underscored, and bold type indicates items that form a problem for the hypothesis that length explains the preference to use a linking element *en*. Frequency: the number of occurrences of the basic compounds with and without *en* in Google (Dutch pages, December, 2004). Head: the words used as head to lengthen the basic compound. Judgments: rating averages for the basic and lengthened compound pairs. Length: the number of syllables added by the Head, divided by the number of syllables of the Basic compound without the linking element.

<table>
<thead>
<tr>
<th>Basic compound</th>
<th>Frequency</th>
<th>Head</th>
<th>Judgments</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>aandelenstorting 'stock call'</td>
<td>7, 21</td>
<td>+s+bewijs 'confirmation'</td>
<td>4.94, 3.89</td>
<td>.50</td>
</tr>
<tr>
<td>aardappelenoogst 'potato harvest'</td>
<td>12, 6750</td>
<td>+machine 'machine'</td>
<td>1.47, 1.53</td>
<td>.75</td>
</tr>
<tr>
<td>autowegenvignet 'motorway sticker'</td>
<td>160, 16</td>
<td>+kosten 'costs'</td>
<td>3.67, 3.19</td>
<td>.40</td>
</tr>
<tr>
<td>bloemenbak 'flower box'</td>
<td>551, 9460</td>
<td>+ontwerp 'design'</td>
<td>2.19, 1.92</td>
<td>1.00</td>
</tr>
<tr>
<td>boekenbinders 'book binder'</td>
<td>36, 684</td>
<td>+opleiding 'training'</td>
<td>2.86, 1.53</td>
<td>1.00</td>
</tr>
<tr>
<td>brievenhoofd 'letter heading'</td>
<td>110, 5440</td>
<td>+voorschrift 'prescription'</td>
<td>2.78, 2.61</td>
<td>1.00</td>
</tr>
<tr>
<td>cellentekort 'cell shortage'</td>
<td>3900, 129</td>
<td>+probleem 'problem'</td>
<td>4.69, 4.89</td>
<td>.67</td>
</tr>
<tr>
<td>dierenverzorgers 'animal attendents'</td>
<td>952, 1100</td>
<td>+handboek 'handbook'</td>
<td>5.64, 4.86</td>
<td>.50</td>
</tr>
<tr>
<td>doelpuntenmakers 'goal makers'</td>
<td>4330, 372</td>
<td>+ranglijst 'priority'</td>
<td>4.14, 4.5</td>
<td>.50</td>
</tr>
<tr>
<td>drankengroothandel 'liquor wholesale'</td>
<td>1170, 34</td>
<td>+vergunning 'license'</td>
<td>2.92, 2.69</td>
<td>.75</td>
</tr>
<tr>
<td>druktoetsentelefoon 'button telephone'</td>
<td>14, 216</td>
<td>+winkel 'shop'</td>
<td>2.92, 3.58</td>
<td>.40</td>
</tr>
<tr>
<td>frisdrankenindustrie 'softdrink industry'</td>
<td>428, 497</td>
<td>+directeur 'manager'</td>
<td>2.67, 2.89</td>
<td>.60</td>
</tr>
<tr>
<td>gordijnemarkt 'curtain market'</td>
<td>12, 7</td>
<td>+standplaats 'stand'</td>
<td>5.33, 4.86</td>
<td>.67</td>
</tr>
<tr>
<td>Basic compound</td>
<td>Frequency</td>
<td>Head</td>
<td>Judgments</td>
<td>Length</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------</td>
<td>------</td>
<td>-----------</td>
<td>--------</td>
</tr>
<tr>
<td>handdoekenhanger ‘towel rack’</td>
<td>11, 31</td>
<td>+ontwerp ‘design’</td>
<td>3.5, 3.47</td>
<td>.50</td>
</tr>
<tr>
<td>handtekeningenstempel ‘signature stamp’</td>
<td>11, 26</td>
<td>+inktdoos ‘inking pad’</td>
<td>3.64, 3.5</td>
<td>.33</td>
</tr>
<tr>
<td>huisdierenvoer ‘pet-animal food’</td>
<td>71, 62</td>
<td>+verpakking ‘wrapping’</td>
<td>3.94, 3.58</td>
<td>1.00</td>
</tr>
<tr>
<td>huurprijzenwet ‘rent law’</td>
<td>910, 11</td>
<td>+ellende ‘trouble’</td>
<td>4.33, 4.61</td>
<td>1.00</td>
</tr>
<tr>
<td>kantorenaanbod ‘office supply’</td>
<td>121, 264</td>
<td>+toename ‘increase’</td>
<td>2.97, 2.64</td>
<td>.75</td>
</tr>
<tr>
<td>klassenonderwijzer ‘classroom teacher’</td>
<td>32, 25</td>
<td>+invaller ‘substitute’</td>
<td>5.56, 5.22</td>
<td>.60</td>
</tr>
<tr>
<td>klerenhanger ‘clothes hanger’</td>
<td>484, 3170</td>
<td>+verzameling ‘collection’</td>
<td>3.17, 2.61</td>
<td>1.25</td>
</tr>
<tr>
<td>maaltijdenservice ‘meal service’</td>
<td>479, 4070</td>
<td>+busje ‘bus’</td>
<td>2.58, 3.25</td>
<td>.40</td>
</tr>
<tr>
<td>normenverval ‘standards decline’</td>
<td>18, 188</td>
<td>+aanpak ‘policy’</td>
<td>3.78, 4.58</td>
<td>.67</td>
</tr>
<tr>
<td>pannendeksel ‘pan lid’</td>
<td>114, 42</td>
<td>+handvat ‘handle’</td>
<td>5.67, 5.19</td>
<td>.67</td>
</tr>
<tr>
<td>pannenkoekenmix ‘pancake mix’</td>
<td>489, 293</td>
<td>+verpakking ‘wrapping’</td>
<td>4.97, 4.64</td>
<td>.75</td>
</tr>
<tr>
<td>paspoortenfraude ‘passport fraud’</td>
<td>33, 124</td>
<td>+procedure ‘procedure’</td>
<td>1.78, 2.11</td>
<td>1.00</td>
</tr>
<tr>
<td>richtlijnenstelsel ‘guidelines system’</td>
<td>17, 8</td>
<td>+structuur ‘structure’</td>
<td>5.11, 4.67</td>
<td>.50</td>
</tr>
<tr>
<td>visserenvoer ‘fish bait’</td>
<td>968, 5220</td>
<td>+verkoper ‘seller’</td>
<td>2.72, 2.47</td>
<td>1.50</td>
</tr>
<tr>
<td>werktijdenschema ‘working-times scheme’</td>
<td>17, 10</td>
<td>+papier ‘paper’</td>
<td>4.53, 4.28</td>
<td>.50</td>
</tr>
<tr>
<td>werktuigenberging ‘tools storeroom’</td>
<td>793, 63</td>
<td>+s+plaats ‘location’</td>
<td>2.64, 2.25</td>
<td>.25</td>
</tr>
<tr>
<td>woningenbestand ‘housing stock’</td>
<td>640, 6230</td>
<td>+computer ‘computer’</td>
<td>3.47, 3.78</td>
<td>.75</td>
</tr>
</tbody>
</table>
Appendix 3

Metrical grids and rhythmic structures of the test items (Study 3 and 4), with comments added where necessary. The construction of metrical grids is regular to a large extent, see for rules and exceptions in Dutch word stress: Kager (1989), Neijt, and van Heuven (1992), Trommelen and Zonneveld (1999), Neijt (2005), Gussenhoven (to appear), and Ernestus and Neijt (to appear); the rules for compound stress are presented in Visch (1989). In this appendix we used labels for the smaller domains (syllable $\sigma$ and foot $\Sigma$) embedded in larger domains (prosodic word $\omega$ and compound word $W$). These are irrelevant to the rhythmic interpretation of the structures, though they are relevant for the assignment of stress (initial stress for feet and compound words, final stress for words) and they are assumed to be relevant for the application of phonological rules that are of no concern to us here.

In (i) the test items are presented for which addition of linking $en$ improves the rhythmic structure in basic compounds, but not in lengthened ones. In (ii) the examples are gathered in which the choice of linking $en$ is neutral with respect to Perfect Rhythm. Addition of linking $en$ has no effect at all on the rhythmic structure or a similar effect in basic and lengthened compound pairs. For instance, in (ii−a) all four forms are rhythmically perfect, and in (ii−b), the choice of linking $en$ leads to a similar rhythmic difference in both compound pairs: two lapses in the form without $en$, or three lapses in the form with $en$. On the basis of rhythm, we expect to find no differences between basic and lengthened compound pairs in such cases. In (iii) three remaining instances are given in which the use of linking $en$ leads to an improved or impoverished rhythmic pattern in one of the compound pairs by violation of Condition (6c) only. In Study 4, (i) is compared with (ii) and (iii).

(i) Improvement in basic pairs

a.

\[
\begin{align*}
(* & )_W \\
(* & )_W ( & *)_\omega \\
(* & )_\omega ( & *)_\omega \\
(* & )_\Sigma ( & *)_\Sigma ( & *)_\Sigma ( & *)_\Sigma \\
(* & )_\sigma ( & *)_\sigma ( & *)_\sigma ( & *)_\sigma ( & *)_\sigma \\
\text{bloem (en) bak ont werp} \\
#bloembäk/bloemênbäk \\
bloembäköntwärp/bloemênbäköntwärp \\
‘flowerbox design’
\end{align*}
\]
b. \[ (*) \] 
\[ (*) \_W ( *) \_W \]
\[ (*) \_W ( *) \_W \]
\[ (*) \_W ( *) \_W \]
\[ (*) \_E ( *) \_E \]
\[ (*) \_E ( *) \_E \]
\[ (*) \_E ( *) \_E \]
\[ (*) \_E ( *) \_E \]
\[ (*) \_E ( *) \_E \]
\[ (*) \_E ( *) \_E \]

boek (en) bin ders op lei ding
#bøekbïnders/bøekbïnders
bøekbïndersøpleïding/bøekbïndersøpleïding
‘bookbinder training’

c. \[ (*) \] 
\[ (*) \_W ( *) \_W \]
\[ (*) \_W ( *) \_W \]
\[ (*) \_W ( *) \_W \]
\[ (*) \_E ( *) \_E \]
\[ (*) \_E ( *) \_E \]
\[ (*) \_E ( *) \_E \]
\[ (*) \_E ( *) \_E \]

brief (en) hoofd voor schrift
#briefhoofof/brïevënhooofd
brïëfëhoofdvooorschrift/brïevënhoofofdvooorschrift
‘letter heading prescription’

d. \[ (*) \] 
\[ (*) \_W ( *) \_W \]
\[ (*) \_W ( *) \_W \]
\[ (*) \_W ( *) \_W \]
\[ (*) \_E ( *) \_E \]
\[ (*) \_E ( *) \_E \]
\[ (*) \_E ( *) \_E \]
\[ (*) \_E ( *) \_E \]

drank (en) groot han del ver gun ning
#~dränkgrëoðhändë/ dränkǹgروحëhändë
~~dränkgrëoðhändëlëvgùnnìng/ ~dränkǹgروحëhändëlëvgùnnìng
‘liquor wholesale business license’
e. \( \begin{array}{c}
( \ast \\ \ast \\
( \ast \\
( \ast \\
( \ast \\
( \ast \\
( \ast \\
\end{array} \)_{\omega}
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
\)
(ii) No difference, or a similar effect in both pairs

a.  
\[
\begin{align*}
\text{pan (en) & dek sel hand vat} \\
\text{påndéksël & pännendéksël} \\
\text{påndéksélhåndvat & pännendéksélhåndvat} \\
\text{‘pan-lid handle’} \\
\text{vis (en) & voer ver ko per} \\
\text{vīsvōer & vīssēnvōer} \\
\text{vīsvōervērkōpēr & vīssēnvōervērkōpēr} \\
\text{‘fish bait seller’}
\end{align*}
\]

\[
\begin{align*}
\text{aan deel (en) & stor tings be wijs} \\
\text{āandēelstōrtīng & āandēlēnstōrtīng} \\
\text{āandēelstōrtīngsbēwijs & āandēlēnstōrtīngsbēwijs} \\
\text{‘stock-call confirmation’}
\end{align*}
\]
b. \( (*) \) \\
\( (*) \) \\
\( (*) \) \\
\( (*) \) \\
\( (*) \) \\
\( *)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma \\
\( \text{aard appel (en) oogst ma chi ne}^3 \) \\
\( \text{‘potato-harvest machine’} \)

c. \( (*) \) \\
\( (*) \) \\
\( (*) \) \\
\( (*) \) \\
\( *)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma \\
\( \text{cel (en) te kort probleem} \) \\
\( \text{‘cell-shortage problem’} \)

d. \( (*) \) \\
\( (*) \) \\
\( (*) \) \\
\( (*) \) \\
\( *)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma \\
\( \text{dier (en) ver zorgers handboek} \) \\
\( \text{‘animal-attendant handbook’} \)

---

\(^3\) Here, a left-branching compound structure is used, reflecting the meaning ‘machine for potato harvesting’. A right-branching structure would reflect a slightly different meaning: ‘harvesting machine for potatoes’. The position of peaks and valleys is similar, however, illustrating the fact that rhythm does not distinguish these alternative meanings.
e. \( (* \quad \quad ) \rlap{\_\_\_\_}\) \\
\( (* \quad \quad ) \rlap{\_\_\_\_}\) \( (* \quad \quad ) \rlap{\_\_\_\_}\) \\
\( (* \quad \quad ) \rlap{\_\_\_\_}\) \( (* \quad \quad ) \rlap{\_\_\_\_}\) \\
\( (* \quad \quad ) \rlap{\_\_\_\_}\) \( (* \quad \quad ) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \\
doel punt (en) ma kers rang lijst
dōelpŭntmākĕrs/dōelpūntmākĕrs
dōelpūntmākĕrsranglijst/dōelpūntmākĕrsranglijst
'goal maker priority list'

f. \( (* \quad \quad ) \rlap{\_\_\_\_}\) \\
\( (* \quad \quad ) \rlap{\_\_\_\_}\) \( (\quad (*) \rlap{\_\_\_\_}\) \\
\( (* \quad \quad ) \rlap{\_\_\_\_}\) \( (\quad (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (\quad (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \\
fris drank (en) in du strie di rec teur
``frīsdrānkīndūsstrē/frīsdrānkēnīndūsstrē``
``frīsdrānkīndūsstrēdīrēkētēur/frīsdrānkēnīndūsstrēdīrēkētēur``
'soft drink industry manager'

g. \( (* \quad \quad ) \rlap{\_\_\_\_}\) \\
\( (* \quad \quad ) \rlap{\_\_\_\_}\) \( (\quad (*) \rlap{\_\_\_\_}\) \\
\( (* \quad \quad ) \rlap{\_\_\_\_}\) \( (\quad (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (\quad (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \\
\( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \( (*) \rlap{\_\_\_\_}\) \\
hand doek (en) han ger ont werp
``händddōekhängēr/händddōekēnāngēr``
``händddōekhängērōntwērp/händddōekēnāngērōntwērp``
'towel−rack design'
h. \[ (* ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ ( * ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ ( * ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ ( * ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ ( * ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ ( * ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ ( * ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ ( * ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ ( * ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ ( * ) \_ ( * ) \_ \]

hand te ke ning (en) stem pel inkt doos
~ händtēkēningestingstempēl/händtēkēningēnstēmpēl
~ händtēkēningstempēlēnktdōos/händtēkēningēnstēmpēlēnktdōos
‘signature stamp inking pad’

i. \[ (* ) \_ \]
\[ (* ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ \]

huis dier (en) voer ver pak king
hūisdīervōer/hūisdīervōer
hūisdīervōervērpkīng/hūisdīervōervērpkīng
‘pet-animal food’

j. \[ (* ) \_ \]
\[ (* ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ \]
\[ (* ) \_ ( * ) \_ \]

huur prijs (en) wet el len de

---

4 The word tekeningen ‘signatures’ has main stress on the preantepenultimate syllable. We assume two feet in this word, with main stress on the first one. Alternative analyses may result in different stretches of lapses in the forms compared, but not in violations of the crucial Condition (6a) Recurrence.
‘rent-law trouble’

klas (en) on der wij zer in val ler

‘classroom teacher substitute’

maal tijd (en) ser vice bus je

‘meal service bus’

---

5 The position of stress is not predictable in verbal compounds such as onderwijzen ‘to teach’ (lit. ‘under + direct/show’) and invallen ‘to substitute’ (lit. ‘in + fall’), which form the bases for the agentive nouns onderwijzer and invaller in this example.

6 As in English, service is a word with two syllables.
m. \( (* \quad )_W \)
\( (* \quad )_W ( \quad *)_\omega \)
\( (* \quad )_W (\quad *)_\omega ( \quad *)_\omega \)
\( (* \quad )_\omega (\quad *)_\omega (\quad *)_\omega ( \quad *)_\omega \)
\( (* \quad )_\Sigma (\quad *)_\Sigma (\quad *)_\Sigma (\quad *)_\Sigma (\quad *)_\Sigma \)
\( (* \quad )_\sigma (\quad *)_\sigma (\quad *)_\sigma (\quad *)_\sigma (\quad *)_\sigma \)

pan nen koek (en) mix ver pak king
pännenköökkmix/pännenköökkmix
pännenköökkmixverpákking/pännenköökkmixværpaṅkking
‘pancake mix wrapping’

n. \( (* \quad )_W \)
\( (* \quad )_W ( \quad *)_\omega \)
\( (* \quad )_\omega ( \quad *)_\omega ( \quad *)_\omega \)
\( (* \quad )_\omega (\quad *)_\omega ( \quad *)_\omega ( \quad *)_\omega \)
\( (* \quad )_\Sigma (\quad *)_\Sigma (\quad *)_\Sigma (\quad *)_\Sigma (\quad *)_\Sigma \)
\( (* \quad )_\sigma (\quad *)_\sigma (\quad *)_\sigma (\quad *)_\sigma (\quad *)_\sigma \)

pas poort (en) frau de pro ce du re
päspsörtfräudě/päspsörtfräudě
päspsörtfräuděpröcěduře/päspsörtfräuděpröcěduře
‘passport fraud procedure’

o. \( (* \quad )_W \)
\( (* \quad )_W ( \quad *)_\omega \)
\( (* \quad )_\omega ( \quad *)_\omega ( \quad *)_\omega \)
\( (* \quad )_\omega (\quad *)_\omega ( \quad *)_\omega ( \quad *)_\omega \)
\( (* \quad )_\Sigma (\quad *)_\Sigma (\quad *)_\Sigma (\quad *)_\Sigma (\quad *)_\Sigma \)
\( (* \quad )_\sigma (\quad *)_\sigma (\quad *)_\sigma (\quad *)_\sigma (\quad *)_\sigma \)

richt liijn (en) stel sel struc tuur
richtlijnstelsēl/richtlijněnstēlsēl
richtlijnstelsēlstruktūr/richtlijněnstēlsēlstruktūr
‘guidelines system structure’
Because *berging* and *bergingsplaats* are used interchangeably in Dutch, the right-branching analysis of this word is most plausible.

---

Language and Speech
A. Neijt, R. Schreuder 565
(iii) Remaining cases

a. \((**)_W (**)_\omega (***)_\omega (****)_\omega\)

\((**)_W (*)_\omega (**)_\omega (***)_\omega \)

\((**)_\omega (**)_\omega (**)_\omega (**)_\omega \)

\((**)_\Sigma (**)_\Sigma (**)_\Sigma (**)_\Sigma \)

\((**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma \)

au to weg (en) vig net kos ten
äutowégvignēt/äutowēgēnvignēt
äutowēgvignētköstēn/“äutowēgēnvignētköstēn
‘motorway sticker costs’

b. \((**)_W (**)_\omega \)

\((**)_W (**)_\omega \)

\((**)_\omega (**)_\omega (***)_\omega \)

\((**)_\Sigma (**)_\Sigma (**)_\Sigma (**)_\Sigma \)

\((**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma \)

druk toets (en) te le soon win kel
~drüktōetsēntēlion / drüktōetsēntēlion
~drüktōetsēntēlionwinkēl / drüktōetsēntēlionwinkēl
‘button telephone shop’

c. \((**)_W (**)_W \)

\((**)_\omega (**)_\omega (***)_\omega (**)_\omega \)

\((**)_\Sigma (**)_\Sigma (**)_\Sigma (**)_\Sigma \)

\((**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma (**)_\sigma \)

norm (en) ver val aan pak
nörmvērvāl / nörmēnvērvāl
nörmvērvālāanpāk / ~nörmēnvērvālāanpāk
‘standards decline policy’