Word length and the location of primary word stress in Dutch, German, and English*

MIRJAM ERNESTUS AND ANNEKE NEIJT

Abstract

This study addresses the extent to which the location of primary stress in Dutch, German, and English monomorphemic words is affected by the syllables preceding the three final syllables. We present analyses of the monomorphemic words in the CELEX lexical database, which showed that penultimate primary stress is less frequent in Dutch and English trisyllabic than quadrisyllabic words. In addition, we discuss paper-and-pencil experiments in which native speakers assigned primary stress to pseudowords. These experiments provided evidence that in all three languages penultimate stress is more likely in quadrisyllabic than in trisyllabic words. We explain this length effect with the preferences in these languages for word-initial stress and for alternating patterns of stressed and unstressed syllables. The experimental data also showed important intra- and interspeaker variation, and they thus form a challenging test case for theories of language variation.[1]

1. Introduction

Most analyses of primary word stress in Dutch, German, and English assume that the location of primary stress is independent of the position of secondary stresses to the left of the primary stress. Thus, in all accounts of primary stress that are rule-based or grounded in the Principles and Parameters framework, the parsing of syllables into feet is from right to left, and primary stress is assigned to the right-most foot. Since the word-final syllable is left extrametric under some conditions, at least at the stage where primary stress is assigned, this results in primary stress on the antepenultimate, the penultimate, or the final syllable (e.g., Hayes 1980; van der Hulst 1984; Kager 1989; Venneman 1990). The footing of the syllables to the left of the three-syllable word-final window thus plays no role in primary stress assignment, as is explicitly stated in van der

Linguistics 46–3 (2008), 507–540 0024–3949/08/0046–0507
DOI 10.1515/LING.2008.017

Hulst and Kooij (1992), who claimed that primary stress assignment and secondary stress assignment are separate algorithms ("main stress first approach").

Also in most analyses framed within Optimality Theory, the position of secondary stress does not affect the position of primary stress (e.g., Nouveau 1994; Féry 1998; Hammond 1999; Pater 2000). The analyses adopt constraints stating that every prosodic word ends in a foot and that primary stress is on the head of this foot. These constraints are ranked high, and may only be violated in order to satisfy constraints on the well-formedness of feet (such as, feet are binary, feet are left-headed, heavy syllables cannot be in the dependent positions of feet) and by constraints forbidding stress on the word-final syllable. Importantly, the constraint forcing all syllables to be parsed into feet is low-ranked. As a consequence, the assumed constraint hierarchies force well-formed feet at the right-edge of the word. The preceding syllables are parsed into feet only if the resulting feet are well-formed.

Germanic languages, as many other languages, tend to have word-initial (secondary) stress (e.g., Halle and Kenstowicz 1991; Booij 1995; Féry 1998; Trommelen and Zonneveld 1999a: 484; Pater 2000; Zonneveld and Nouveau 2004). Speakers transfer this preference for initial stress in existing words to pseudowords (Baker and Smith 1976), and listeners take advantage of it in speech segmentation (Cutler and Norris 1988) and in phonetic judgments (Slootweg and Rietveld 1989). The preference for initial stresses may overrule other constraints. For instance, the Dutch word *ˌJerosoˌlimi ˈtani* 'friars of the order of Jerusalem' bears word-initial secondary stress, even though this results in a sequence of two unstressed syllables (Zonneveld and Nouveau 2004: 389), and Germanic languages, like many other languages, prefer alternating patterns of stressed and unstressed syllables (e.g., Prince 1983; Selkirk 1984; Kager 1989; van der Hulst and Kooij 1992; Hung 1994; Wiese 1996). The preference for initial stress is not satisfied, however, if it leads to unwell-formed feet or stress clashes. More importantly, it is claimed that it would not affect the position of primary stress. Thus, Pater (2000) and Zonneveld and Nouveau (2004) state that the constraint forcing words to start with a stressed syllable is satisfied only when this can be done at no expense of the constraint Align (Prosodic Word, Right, Head of Prosodic Word, Right), which forces the head of the prosodic word to be at the word-final syllable.

In the present study, we questioned the independence of primary stress of the footing of the preceding syllables. More specifically, we investigated the role of the preference for word-initial stress in combination with the preference for alternating patterns of stressed and unstressed syllables. For this we compared tri- and quadrisyllabic existing words and pseudowords in Dutch, German, and English. If the location of primary stress is determined only on the basis of the three final syllables, we may expect that the number of syllables in the word is irrelevant. Trisyllabic and quadrisyllabic words ending in the same three syllables would have primary stress on the same syllable (counting from the right edge). If, in contrast, the footing of all syllables in the word would affect the location of primary stress, we may expect a difference between trisyllabic and quadrisyllabic words. The preference for stressed initial syllables in combination with the preference for alternating patterns of stressed and unstressed syllables may favor penultimate primary stress in quadrisyllabic words (ˌσσˈ σσ, with ˈ indicating primary and ˌ indicating secondary word stress). In contrast, trisyllabic words with secondary stress on the initial syllable and primary stress on the penultimate syllable would not show a perfect alternation of stressed and unstressed syllables (ˌσˈ σσ), but a stress clash. Such words are predicted to favor initial, antepenultimate, primary stress.

In words ending in syllables of certain phonological structures, however, we predict that the preference for stressed word-initial syllables may be overruled by other (language-specific) preferences, such as the constraint that forces heavy syllables to be the head of feet (e.g., Neijt and Zonneveld 1982; Prince and Smolensky 1993; Zonneveld and Nouveau 2004). Such trisyllabic words may not have initial stress, and we predict that the corresponding quadrisyllabic words bear primary stress on the same syllable (counting from the right edge) as their trisyllabic counterparts. We thus expect a difference between tri- and quadrisyllabic words especially for word types of which the trisyllabic words tend to have initial, antepenultimate, stress.

The literature contains several indications that the location of primary stress may indeed be affected by the number of syllables in the word. Don and Zonneveld (1988: 20) observed that 13 out of the 21 tri- and quadrisyllabic Dutch words with primary stress on the penultimate open syllable followed by a closed final syllable (ˈCVCVC) are quadrisyllabic, even though quadrisyllabic words are much less frequent than trisyllabic words in Dutch. Don and Zonneveld speculated that this may be because penultimate stress in quadrisyllabic words allows a foot at the left word-edge, which results in full footing and a eurythmic pattern of stressed and unstressed syllables (ˌσσˈ σσ).

Additional support for the effect of word length on the position of primary stress in Dutch comes from Nouveau (1994: 105). She asked twenty speakers of Dutch to read aloud a number of pseudowords, including trisyllabic *merotak* and *dapiton*, and quadrisyllabic *monitaron.* These three words end in syllables of the same phonological structure (open - open - closed), but *monitaron* differs from the two others in the number of syllables. Nouveau found that whereas the two trisyllabic words received penultimate stress in only 10% and 5% of the realizations, respectively (most speakers preferred initial, antepenultimate, stress), the quadrisyllabic word was assigned penultimate stress in 40% of cases ($\chi^2_1 = 3.33$, $p = 0.07$; $\chi^2_1 = 5.16$, $p < 0.05$, respectively. All statistical tests reported in this article are two-tailed). This suggests that word length affects primary stress assignment.

Janßen (2003) carried out a similar experiment in German. She asked 28 native speakers of German to produce a large number of pseudowords. Four word types, as characterized by the phonological structure of their final syllables, were represented by both trisyllabic and quadrisyllabic pseudowords. Three out of these four word types showed a numerically lower percentage of penultimate stress among the tri- than among the quadrisyllabic words. This difference is significant for the words ending in a closed antepenultimate, an open penultimate, and a closed final syllable ($\chi^2_1 = 4.18$, $p < 0.05$). The trisyllabic words (9 items) carried penultimate stress in 23% (48) of the tokens and the quadrisyllabic word in 44% (11) of its tokens. This suggests that also in German word length may affect the position of primary stress.

In the present study, we first investigated the role of word length on the location of primary stress in existing words of Dutch, German, and English (Section 2). We then carried out paper-and-pencil experiments in which we investigated the effect of word length in pseudowords of these languages (Section 3). The experiments showed variation between pseudowords, within, and between speakers, which we discuss in Section 4. Finally, we summarize and discuss our results in Section 5.

2. Lexical statistics

2.1. *General method*

The CELEX lexical database (Baayen et al. 1995) contains phonological and morphological descriptions of thousands of Dutch, German, and English words. We collected all trisyllabic and quadrisyllabic words that are marked as "monomorphemic", which do not contain productive or semiproductive affixes. We then investigated whether the location of primary stress is affected by the number of syllables in the word. In addition to word length, we considered as independent variables the factors that have repeatedly been identified as relevant to the location of primary stress, that is, the phonological structure of the three final syllables as well as word class for English.

2.2. *Dutch*

In Dutch, primary stress tends to fall on the penultimate syllable (e.g. [1a], [1b]), except in four types of words (e.g., Neijt and Zonneveld 1982; van der Hulst 1984; Don and Zonneveld 1988; Kager 1989; Trommelen 1991; Nouveau 1994; Trommelen and Zonneveld 1999b; Zonneveld and Nouveau 2004; Gussenhoven i.p.). Primary stress necessarily falls on the final syllable of monosyllabic words. It also falls on the final syllable of words of which this final syllable (a) ends in a tense vowel and at least one coda consonant, (b) in a lax vowel and at least two coda consonants, or (c) contains a diphthong (i.e., super-heavy syllables; as in [1c], [1d], [1e]). Furthermore, primary stress tends to fall on the antepenultimate syllable of words with an open penultimate syllable and a final syllable ending in a lax vowel and one consonant (e.g., [1f], [1g]). Finally, syllables with schwa never bear stress.

(1) Examples for Dutch stress patterns (syllable boundaries are indicated by ".")

a. a.'man.del 'almond'
b. an.'dij.vie 'endive'
c. a.mal.'gaan 'amalgan'
d. ar.gu.'ment 'argument'
e. a.ke.'lei 'columbine'
f. 'ma.ra.thon 'marathon'
g. 'car.na.val 'carnival'

These generalizations indicate that for Dutch, the distinction between syllables with schwa, open syllables (ending in a tense vowel), closed syllables (ending in a lax vowel and one consonant), and super-heavy syllables (ending in a tense vowel and at least one consonant, a lax vowel and at least two consonants, or containing a diphthong) is relevant.

We collected all monomorphemic words consisting of three or four syllables from the Dutch part of the CELEX lexical database. The resulting data set contained 837 words (out of the 9861 monomorphemic words in CELEX), most of them non-native. We classified the location of primary stress in these words (antepenultimate, penultimate, or final) as given in CELEX, and the final three syllables as containing a schwa, as open, as closed, or as super heavy, along the definitions given above. Note that these definitions imply that tense vowels do not appear in syllables of the

Table 1. *The numbers of words with antepenultimate, penultimate, and final primary stress in the Dutch database broken down by the number of syllables (NS), and the phonological structures of the penultimate and final syllable*

Word type			Location of primary stress		
NS	Penult	Final	Antepenult	Penult	Final
3	schwa	schwa	1	0	0
3	super heavy	schwa	0	1	0
3	closed	schwa	0	71	0
4	closed	schwa	0	15	0
3	open	schwa	6	135	0
4	open	schwa	0	43	1
3	schwa	open	8	0	8
3	super heavy	open	0	1	0
3	closed	open	2	41	6
4	closed	open	0	12	0
3	open	open	63	74	24
4	open	open	18	21	2
3	schwa	closed	8	0	9
3	closed	closed	2	2	2
4	closed	closed	0	1	0
3	open	closed	54	7	34
4	open	closed	1	8	5
3	schwa	super heavy	9	0	39
3	closed	super heavy	0	1	7
4	closed	super heavy	0	0	2
3	open	super heavy	5	0	82
4	open	super heavy	0	0	3
4	super heavy	closed	2	0	0
4	super heavy	super heavy	1	0	0

category "closed", since syllables with tense vowels and a coda consonant are super heavy. Similarly, syllables with diphthongs do not occur among closed or open syllables, since syllables with diphthongs are super heavy. Finally, lax vowels do not appear in open syllables, since lax vowels are followed by at least one consonant in word-final syllables, and they are presumably followed by at least an ambisyllabic consonant in word-medial syllables (see the literature starting with van der Hulst 1985). Furthermore, we classified the words as consisting of three or four syllables. Table 1 gives the numbers of words with antepenultimate, penultimate, and final primary stress for the attested types of words, as characterized by the phonological structures of the final two syllables and word length.

We analyzed the data set with a Classification and Regression Tree Analysis (CART, Breiman et al. 1984). This technique divided the data set into groups, with words that are identical in their length and in the

Table 2. *The CART groups (G) of Dutch existing words, the number of words (NW) in each group, and the percentages of words with primary stress on the antepenultimate (Apenult), the penultimate, and the final syllable. The groups are characterized by the number of syllables (NSyll) in the words and the phonological structure of the penultimate and final syllable. A " — " indicates that the number of syllables or the phonological structures of the penultimate syllables is irrelevant*

Group			Characterization		Location of primary stress		
G	NW	Nsyll	Penult	Final	Apenult (%)	Penult (%)	Final (%)
1	273	—	—	schwa	2.6	97.1	0.4
2	16	—	schwa	open	50.0	0.0	50.0
3	202	—	open	open	40.1	47.0	12.9
4	62	—	closed, super heavy	open	3.2	87.1	9.7
5	118	three	—	closed	54.2	7.6	38.1
6	17	four	—	closed	17.6	52.9	29.4
7	149	—	—	super heavy	10.1	0.7	89.2

phonological structure of their three final syllables (henceforth: of the same type) being grouped together. The types of words in a group are similar in their preference for the location of primary stress, and the characterization of the groups therefore provides information on which factors affect the location of primary stress.

Table 2 lists the seven groups created by CART, as well as the numbers of words, and the percentages of words with antepenultimate, penultimate, and final primary stress in these groups. The phonological structure of the antepenultimate syllable is not part of the characterization of any group, which is in line with the observation in the literature that only the phonological structures of the final and the penultimate syllables affect the location of primary stress in Dutch (see above). The percentages of words with antepenultimate, penultimate, and final stress in the groups support the observation that the unmarked location of primary stress is on the penultimate syllable (Groups 1, 3, 4, 6). If this syllable contains a schwa, and consequently cannot bear stress, primary stress falls equally often on the antepenultimate and on the final syllable (Group 2). Also, as described in the literature, primary stress tends to fall on the final syllable if this syllable is super heavy (Group 7).

The CART analysis does not support the observation that for primary stress to fall on the antepenultimate syllable, the final syllable has to be closed and the penultimate syllable to be open. For words ending in closed syllables, the phonological structure of the penultimate syllable seems to be irrelevant (Groups 5 and 6), probably because nearly all such words in the data set contain open penultimate syllables (128 out of

135 words). The important predictor for words of this type appears to be the number of syllables (compare Groups 5 and 6): The percentage of words with penultimate stress is larger among the quadrisyllabic (52.9%) than the trisyllabic (7.6%) words. A chi-squared test confirms that we are not observing a random pattern ($\chi^2 = 8.58$, $p < 0.01$). Two native speakers, who were blind to the purpose of this study, ascertained that these quadrisyllabic words with penultimate primary stress all have secondary initial stress, as we would expect if the word-initial stress drives the length effect.

Importantly, we observe exactly the same patterns of results, if we restrict our data set to words that have at least two non-reduced vowels in their final three syllables (these results are not included in the table). Only in these words, the location of primary stress is not completely fixed by the phonological structure of the final three syllables. We conclude that word length affects the location of primary stress in Dutch.

We observed an effect of word length only for the words ending in a closed syllable. The trisyllabic words of this type have initial stress in 54.2% of cases. We observed no length effect for the words ending in two open syllables, which is the group of words with the second highest percentage ($63/(63 + 74 + 24) * 100\% = 39.1\%$, cf. Table 1) of trisyllabic words with word-initial stress (apart from the trisyllabic words without quadrisyllabic counterparts in the data set). The majority of these trisyllabic words ending in two open syllables (at least $100\% - 39.1\% = 60.9\%$) seem to fulfill other constraints than the preference for initial stress, which may block the length effect. Remark, however, that the words do numerically show the predicted effect of word length: The percentage of penultimate stress is slightly higher among the quadrisyllabic ($21/(18+21+2) * 100\% = 51.2\%$, cf. Table 1) than among the trisyllabic words ($74/(63+74+24) * 100\% = 46.0\%$) ending in two open syllables.

2.3. *German*

The Dutch and German stress systems are similar in that also in German primary stress tends to fall on super-heavy final syllables ([2a] and [2b]), and on the penultimate syllable of words without super-heavy final syllables ([2c] and [2d]). According to Wiese (1996), Jessen (1999), and Janßen (2003), antepenultimate stress, instead of penultimate stress, primarily occurs in words with open penultimate syllables, while Féry (1998) stated that antepenultimate stress can be found among all words ending in nonsuper-heavy syllables (see [2e] and [2f]). Féry (1998) also explicitly stated that there is no difference with respect to the location of primary stress between trisyllabic and quadrisyllabic words, except for words that sound as compounds or are grammatical terms.

(2) Examples for German stress patterns (syllable boundaries are indicated by ".")

a.	A.pa.'rat	'apparatus'
b.	Ma.nus.'kript	'manuscript'
c.	In.'spek.tor	'inspector"
d.	A.'re.na	'arena'
e.	'Pa.pri.ka	'pepper'
f.	'Ka.bel.jau	'cod'

We collected all German words that consist of three or four syllables and are coded as "monomorphemic" in the CELEX lexical database. We purged this list manually of the large number of words that are obviously morphologically complex (e.g., *Abwesenheit, Entstalinisierung, Reinmachefrau*). We also excluded from the data set the words with unclear synchronic morphological status and of which the Dutch equivalents do not occur in the Dutch data set (e.g., *Histologie*), such that the final German data set is well comparable to the Dutch data set. The resulting German data set is much smaller (359 words) than the Dutch data set (837 words), which is not surprising since the German part of CELEX contains fewer lemmas (51,728 lemmas) than the Dutch part (124,136 lemmas). For all words in the data set, we classified the location of primary stress (antepenultimate, penultimate, or final) as given in CELEX, and we ascertained the number of syllables. Given the generalizations on German stress in the literature, we classified the three final syllables as containing schwa, as open, as closed, or as super heavy, as we did for Dutch. No syllables with lax vowels are open, since lax vowels are minimally followed by ambisyllabic consonants, as is also indicated in the phonological transcriptions in CELEX. Table 3 shows the number of words with antepenultimate, penultimate, and final primary stress for each attested word type, which is characterized by the phonological structures of the penultimate and final syllables, and the number of syllables in the word.

We analyzed also this data set with a Classification and Regression Tree Analysis, predicting the location of primary stress on the basis of the phonological structures of the three final syllables and word length. Table 4 lists the resulting 6 groups, as well as the number of words, and the percentages of words with antepenultimate, penultimate, and final stress in these groups. The percentages show that penultimate stress is the most frequent stress pattern in German words ending in a syllable with schwa (Groups 1–3), while final stress is most common among

Table 3. *The numbers of words with antepenultimate, penultimate, and final primary stress in the German database broken down by the number of syllables (NS), and the phonological structures of the penultimate and final syllable*

Word type			Location of primary stress		
NS	Penult	Final	Antepenult	Penult	Final
3	schwa	schwa	2	0	0
3	closed	schwa	1	60	0
4	closed	schwa	0	3	0
3	open	schwa	17	89	0
4	open	schwa	5	23	0
3	schwa	closed	1	0	0
3	closed	closed	4	4	2
4	closed	closed	0	1	0
3	open	closed	8	3	4
3	schwa	open	2	0	2
3	closed	open	3	8	5
3	open	open	12	12	7
4	open	open	0	2	2
3	schwa	super heavy	0	0	1
3	closed	super heavy	0	0	21
3	open	super heavy	10	2	36
4	open	super heavy	0	0	2
4	open	closed	2	2	1

Table 4. *The CART groups (G) of German existing words, the numbers of words (NW) in each group, and the percentages of words with primary stress on the antepenultimate (Apenult), the penultimate, and the final syllable. The groups are characterized by the phonological structures of the antepenultimate, the penultimate, and the final syllables. A " — " indicates that the phonological structure of the syllable is irrelevant*

Group			Characterization		Location of primary stress		
G	NW	Apenult	Penult	Final	Apenult (%)	Penult (%)	Final (%)
1	75	schwa, closed	schwa, open	schwa	8.0	92.0	0.0
2	61	open	schwa, open	schwa	29.5	70.5	0.0
3	64	—	closed	schwa	1.6	98.4	0.0
4	87	—	—	open, closed	36.8	36.8	26.4
5	22	—	schwa, closed	super heavy	0.0	0.0	100.0
6	50	—	open	super heavy	20.0	4.0	76.0

words ending in superheavy syllables (Groups 5 and 6). In words ending in open and closed syllables (Group 4), stress is almost equally distributed among the three final syllables, with a slight disadvantage for final stress.

What is important for our research question is that none of the groups created by CART are characterized by word length. If we restrict the data set to words which contain at least two full vowels in their final three syllables, we obtain approximately the same grouping of words. We conclude that CART presents no evidence for an effect of word length on the location of primary stress in German existing words. Given the small number of words in the data set, this non-result is not very surprising.

2.4. *English*

The distribution of primary stress in English is more complex than in Dutch or German. In English, all final syllables with long vowels are stressed. In only few words, however, this final stress is the primary stress: In words of all word classes and of all phonological structures, final stress tends to be secondary, and primary stress is retracted to the penultimate syllable, the antepenultimate syllable (e.g., [3a], [3b]), or, in some exceptional cases, even to the pre-antepenultimate syllable. Final primary stress is mainly found on bisyllabic words (e.g., Chomsky and Halle 1968; Hayes 1980; Kager 1989; Hammond 1999).

For all words ending in syllables with short, instead of long, vowels, word class is an important predictor (e.g., Chomsky and Halle 1968; Hayes 1980, 1982; Kager 1989; Jensen 1993; Hammond 1999). In nouns and some adjectives, primary stress tends to fall on the antepenultimate syllable if the penultimate syllable ends in a short vowel ([3c], [3d]). If not, the penultimate syllable tends to bear stress ([3e], [3f]). In verbs and other adjectives, the final syllable tends to be stressed, if this syllable contains at least two coda consonants. Again, this results in final secondary stress in most words, as primary stress tends to be retracted to the penultimate, antepenultimate, or pre-antepenultimate syllable. If the final syllable does not end in two consonants, stress tends to fall on the penultimate syllable (e.g., [3g], [3h]). Note that these generalizations indicate that penultimate syllables ending in a short vowel and a simple coda do not pattern with syllables only ending in a short vowel in nouns, and that final syllables of the same phonological structure (VC) do not pattern with syllables ending in two consonants in verbs. That is, penultimate VC-syllables in nouns count as heavy, whereas final VC-syllables in verbs count as light.

Finally, syllables without vowel do not count for primary stress assignment (Hammond 1999). Thus, the noun *salamander* ends in a vowel-less syllable in rhotic dialects and has primary stress on its initial syllable, which is as predicted if the final syllable is not taken into account.

(3) Examples for English stress patterns (syllable boundaries are indicated by ".")
   a. sta'lag.mite
   b. 'pa.ra.dise
   c. 'la.by.rinth
   d. A.'me.ri.ca
   e. a.'gen.da
   f. A.ri.'zo.na
   g. a.'sto.nish
   h. de.'ve.lop

We collected all 750 monomorphemic words consisting of three or four syllables from the English part of the CELEX lexical database. The vast majority are nouns. We marked the location of primary stress in these words (antepenultimate, penultimate, or final) as given in CELEX, and classified syllables as super light if their nucleus contained a schwa or a syllabic consonant, as light if they ended in a short vowel, and as heavy if their rimes contained a long vowel or two or more consonants. Syllables ending in a short vowel and a simple coda formed a category by themselves (VC), since the literature showed that they group with either the heavy or the light syllables depending on word class or the position of the syllable in the word. Furthermore, we classified the words as nouns, verbs, or adjectives/adverbs, and as consisting of three or four syllables. Table 5 gives the numbers of nouns with antepenultimate, penultimate, and final stress for the attested word types characterized by the phonological structures of the penultimate and final syllables, and the number of syllables in the word. Table 6 presents the counts for the adjectives, the adverbs, and the verbs.

We analyzed also this data set of monomorphemic words with a Classification and Regression Tree Analysis. Table 7 lists the seven groups representing the CART tree, as well as the numbers of words in these groups, and the percentages of words with antepenultimate, penultimate, and final stress. Groups (1–2) show that primary stress tends to fall on non-light penultimate syllables, while Groups (3–6) show that words with light or super-light penultimate syllables typically bear antepenultimate stress. Group 7 concerns the words with super-light antepenultimate syllables and light or super-light penultimate syllables, which tend to bear primary stress on the penultimate syllable. CART thus wrongly predicts that super-light penultimate syllables may bear primary stress, probably because of scarcity of the data.

Word class and the number of syllables in the word do not play part in the characterizations of the seven word groups. These factors may emerge

Table 5. *The numbers of nouns with antepenultimate, penultimate, and final primary stress in the English database broken down by the number of syllables (NS), and the phonological structures of the penultimate and final syllables ("S.light" stands for "Super light")*

Word type			Location of primary stress		
NS	Penult	Final	Antepenult	Penult	Final
3	VC	VC	1	1	0
3	light	VC	26	6	5
4	light	VC	0	2	0
3	S.light	VC	40	0	7
4	S.light	VC	5	0	1
3	heavy	VC	6	2	1
4	heavy	VC	0	2	0
3	VC	light	0	1	0
3	light	light	3	1	0
4	light	light	2	2	0
3	S.light	light	12	0	0
4	S.light	light	6	0	0
3	heavy	light	0	11	0
3	VC	S.light	6	19	0
4	VC	S.light	1	5	0
3	light	S.light	73	33	0
4	light	S.light	13	3	1
3	S.light	S.light	48	0	0
4	S.light	S.light	6	0	0
3	heavy	S.light	4	60	0
4	heavy	S.light	0	17	0
3	VC	heavy	1	9	0
4	VC	heavy	0	3	0
3	light	heavy	57	8	1
4	light	heavy	10	1	0
3	S.light	heavy	69	0	24
4	S.light	heavy	3	0	0
3	heavy	heavy	8	26	2
4	heavy	heavy	0	9	0

as predictors, however, if we focus on the words that have full vowels in at least two of their three final syllables. Only in these words, the position of primary stress is not completely fixed by the phonological structures of the final syllables. The results of a CART analysis of just these 328 words are given in Table 8. Word class, word length, the phonological structure of the penultimate syllable, and the phonological structure of the final syllable emerge as predictors. The location of primary stress is affected by word class for words with a light penultimate syllable and a final syllable that is light or ends in a short vowel and a consonant (Groups 1 and 2). Stress tends to fall on the penultimate syllable in adverbs/adjectives and

Table 6. *The numbers of adjectives, adverbs, and verbs with antepenultimate, penultimate, and final primary stress in the English database broken down by the number of syllables (NS), and the phonological structures of the penultimate and final syllables ("S.light" stands for "Super light")*

Word type			Location of primary stress		
NS	Penult	Final	Antepenult	Penult	Final
3	light	VC	1	9	0
4	light	VC	0	1	0
3	S.light	VC	2	0	0
3	heavy	VC	1	4	0
4	heavy	VC	0	2	0
3	S.light	light	1	0	0
3	VC	S.light	0	3	0
3	light	S.light	4	5	0
4	light	S.light	2	0	0
3	S.light	S.light	8	0	0
3	heavy	S.light	0	2	0
4	heavy	S.light	0	1	0
3	light	heavy	15	1	0
4	light	heavy	1	0	0
3	S.light	heavy	8	0	2
4	S.light	heavy	1	0	0
3	heavy	heavy	6	4	1
4	heavy	heavy	0	2	0

verbs (Group 1), and on the antepenultimate syllable in nouns (Group 2). This difference between verbs and nouns is expected given the literature. According to our CART-analysis, word class does not play a role for words with other phonological structures, which is possibly due to the small number of verbs in the data set. Thus in words of all classes, primary stress tends to fall on antepenultimate syllables that are followed by light penultimate syllables and heavy final syllables (Group 3), a generalization that according to the literature holds for nouns only.

The CART analysis also showed that word length has an effect on words with heavy or VC penultimate syllables (Groups 4 and 5). All quadrisyllabic words of this type bear primary stress on the penultimate syllable (in combination with secondary stress on the initial syllable, as could be established on the basis of the Van Dale dictionary, 1997). In contrast, penultimate stress can be found in only 57.8% of the trisyllabic words. This difference between the tri- and quadrisyllabic words is statistically significant ($\chi_1^2 = 4.95$, $p < 0.05$), and thus forms evidence of an effect of word length on the location of primary stress in English. The restriction of the effect to words with VC or heavy penultimate syllables is

Table 7. *The CART groups (G) of English existing words, the numbers of words (NW) in these groups, and the percentages of words with primary stress on the antepenultimate (Apenult), the penultimate, and the final syllable. The groups are characterized by the phonological structure of the antepenultimate, the penultimate, and the final syllable. "S.light" stands for super light and a " — " indicates that the phonological structures of the final syllables are irrelevant*

Group			Characterization		Location of primary stress		
G	NW	Antepenult	Penult	Final	Apenult (%)	Penult (%)	Final (%)
1	143	not S.light	VC, heavy	—	23.8	73.4	2.8
2	78	S.light	VC, heavy	—	0.0	100.0	0.0
3	50	not S.light	light	VC, light	62.0	28.0	10.0
4	204	not S.light	light	S.light, heavy	85.8	13.7	0.5
5	161	not S.light	S.light	VC, heavy	79.5	0.0	20.5
6	81	not S.light	S.light	light S.light	100.0	0.0	0.0
7	33	S.light	light, S.light	—	3.0	90.9	6.1

Table 8. *The CART groups (G) of English existing words with full vowels in at least two of the three final syllables, the numbers of words (NW) in these groups, and the percentages of words with primary stress on the antepenultimate (Apenult), the penultimate, and the final syllable. The groups are characterized by word class (verb, noun, adverb/adjective), by the number of syllables (NSyll), and by the phonological structures of the penultimate, and the final syllables. A " — " indicates that the value of the given variable is irrelevant*

Group			Characterization			Location of primary stress		
G	NW	Nsyll	Word class	Penult	Final	Apenult (%)	Penult (%)	Final (%)
1	8	—	A, V	light	VC, light	12.5	87.5	0.0
2	42	—	N	light	VC, light	71.4	16.7	11.9
3	204	—	—	light	heavy	93.2	5.6	1.1
4	10	four	—	VC, heavy	—	0.0	100.0	0.0
5	64	three	—	VC, heavy	—	35.9	57.8	6.3

not unexpected. Some one third of the trisyllabic words of this type bear antepenultimate stress (see Tables 5 and 6). Moreover, the penultimate syllable may bear stress in these words, as it is heavy.

3. Productivity in pseudowords

3.1. *Introduction*

The vast majority of trisyllabic and quadrisyllabic words in the languages under investigation are non-native, and the observed effect of word length in Dutch and English on the location of primary stress may therefore originate from the (contemporary) phonological adaption of these words. If so, the observed effect of word length is part of the contemporary grammars of these languages. It is also possible, however, that the stress patterns of the relevant words reflect older stages of Dutch and English, or that these words adhere to their stress patterns in the source languages (primarily Greek, Latin, and French).

In a series of production experiments, we investigated the contemporary status of the effect of word length on the location of primary stress. Speakers of Dutch, German, and English were presented with orthographic representations of non-existing, but possible, words and were asked to indicate the most likely location for primary stress. If the effect of word length on the location of primary stress is part of the contemporary grammars of the languages, we expect it to affect the location of primary stress in pseudowords.

3.2. *Dutch*

The CART analysis of the Dutch data set (Section 2.2) showed that word length is a predictor for the location of primary stress in existing words ending in closed syllables. We investigated whether also pseudowords of the same phonological structure show the length effect. Moreover, we contrasted these words with words ending in open syllables. Words with final open syllables did not show a statistically significant effect of word length in the lexical analyses. Nevertheless they seem to be good candidates to show the effect, since, after the trisyllabic words ending in an open and a closed syllable, the trisyllabic words ending in open syllables show the highest percentage of initial stress (cf. Table 1). This suggests that the preference for initial stress is also important for this group of words, which may lead to a word length effect.

3.2.1. *Materials.* We created forty quadrisyllabic words obeying the phonotactic restrictions of Dutch. Their three initial syllables consisted of a simple onset and a tense vowel, while their final syllable was closed and contained a lax vowel (e.g. [defapokɛt]). For these words, we created forty trisyllabic counterparts by removing the initial syllable ([fapokɛt]). In addition, we created forty quadrisyllabic and forty trisyllabic versions ending in open syllables, by removing the final consonant of the items in the two sets and changing the final lax vowel into the corresponding tense vowel ([defapoke], [fapoke]). Each word thus had four versions (quadrisyllabic and trisyllabic with a closed final syllable, and quadrisyllabic and trisyllabic with an open final syllable).

In order to have the words in three different orders, we created three master lists, each with three complementary lists. The master lists contained one version of each experimental word. Ten words appeared as quadrisyllabic and ten words as trisyllabic items ending in closed syllables, and ten words appeared as quadrisyllabic and ten words as trisyllabic items ending in open syllables. The three complementary lists for each master list also contained ten tokens of every item type, and together with their master list, these lists contained all experimental items. This procedure resulted in 12 experimental lists. The words in the lists were pseudo randomized such that the items of a given type were equally distributed over the lists. Moreover, a given experimental word occupied the same position in a master list and its three complementary lists.

All items were spelled in the International Phonetic Alphabet, such that their representations unambiguously reflected the intended pronunciations.

3.2.2. *Participants.* Forty-eight native speakers of Dutch participated in the experiment. They were all students of Dutch at the Radboud University Nijmegen, and mastered the International Phonetic Alphabet. They had not yet received any education in stress theory or the stress rules of Dutch.

3.2.3. *Procedure.* Each participant received one of the twelve lists with forty experimental items. Each item was printed three times on the same line. The antepenultimate syllable was underlined in the leftmost token of the word, the penultimate syllable in the middle token, and the final syllable in the right most token (see [4]). The underlining indicated the location of primary stress. Participants were told that the words were nouns, and they were asked to mark for each word the token representing their favorite pronunciation.

(4) First three lines of an experimental list

pesoɣofa pesoɣofa pesoɣofa
ɣedafoke ɣedafoke ɣedafoke
tozadɑp tozadɑp tozadɑp

3.2.4. *Results and Discussion.* Due to a programming error, ten out of the forty trisyllabic items ending in an open final syllable were transcribed with a lax instead of a tense final vowel (e.g., podabɑ instead of podaba). These incorrect transcriptions represented pronunciations that violate the

Table 9. *The absolute numbers and percentages of Dutch responses with antepenultimate, penultimate, and final stress for the quadrisyllabic and trisyllabic items with closed and open final syllables*

Item type		Location of primary stress					
Final syllable	Word length	Antepenult		Penult		Final	
closed	3 syllables	206	57.2%	117	32.5%	37	10.3%
closed	4 syllables	159	44.3%	161	44.8%	39	10.9%
open	3 syllables	113	31.4%	226	62.8%	21	5.9%
open	4 syllables	97	26.7%	249	69.2%	15	4.2%

phonotactic constraints of Dutch. We removed these experimental items and the corresponding items from the final data set, leaving 30 experimental words and 120 experimental items. The resulting experimental items are listed in the appendix. Since one participant failed to indicate the preferred stress pattern for one item, our final data set contained $(48 * 30 - 1 =)1439$ responses. Table 9 lists the absolute numbers and percentages of the responses with antepenultimate, penultimate, and final primary stress broken down for the type of the item (quadrisyllabic/trisyllabic, closed/open final syllable).

We analyzed the penultimate stress responses versus the non-penultimate stress responses with generalized linear mixed models with multivariate normal random effects, using penalized quasi-likelihood (e.g., Schall 1991; Venables and Ripley 2002). We performed two analyses: In one analysis, we considered the participant as random effect, and in the other analysis the experimental item. That is, in one analysis we allowed the participants to differ from each other in how often they assigned penultimate stress as well as in their sensitivity to the independent variables in the analysis. In the other analysis, the experimental items were allowed to differ from each other. In both analyses, our independent variables were Word length (trisyllabic or quadrisyllabic) and the Phonological structure of the final syllable (open or closed).

Both analyses revealed significant main effects for both variables (Word length in the analysis with participant as random variable: $F(1, 1388) = 5.74$, $p < 0.05$; and in the analysis with experimental word as random variable: $F(1, 1406) = 6.84$, $p < 0.01$; Phonological structure of the final syllable in the analysis with participant as random variable: $F(1, 1388) = 59.46$, $p < 0.001$; and in the analysis with experimental word as random variable: $F(1, 1406) = 99.99$, $p < 0.001$). Participants assigned penultimate stress more often to quadrisyllabic words than to trisyllabic words, and more often to words ending in open syllables than to words ending in closed syllables. There was no interaction of

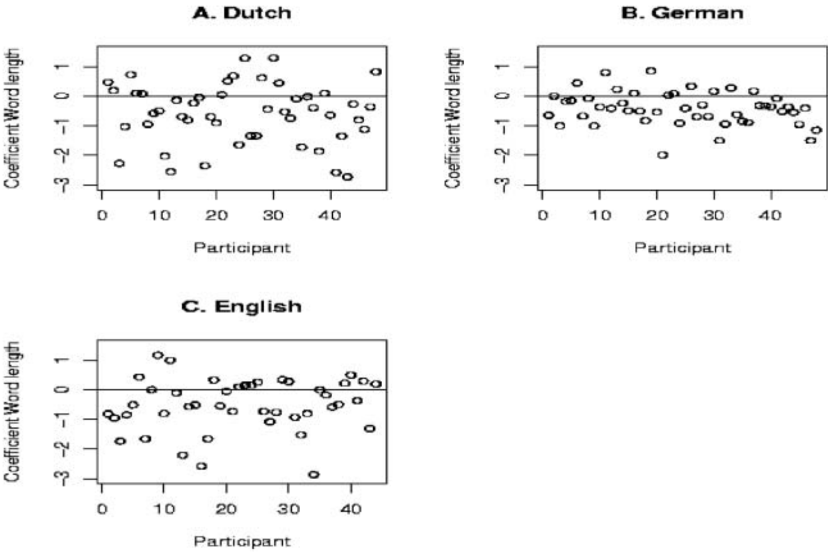

Figure 1. *The coefficients of word length for each of the participants in the Dutch, German, and English paper-and-pencil experiments. A negative coefficient means that the participant tended to assign penultimate primary stress more often to quadrisyllabic than to trisyllabic words*

Word length by Phonological structure of the final syllable ($F < 1$; $F(1, 1406) = 1.50$, $p > 0.1$).

Interestingly, the analysis with participant as random variable showed that the participants differed substantially in their sensitivity to Word Length (log-likelihood ratio = 89.35, $p < 0.001$) and also with respect to the structure of the final syllable (log-likelihood ratio = 26.78, $p < 0.001$). The model suggested for no less than 15 out of 48 participants that they tended to assign penultimate stress less often, instead of more often, to quadrisyllabic words. Figure 1a shows the coefficient for Word length (y-axis) for each participant in the experiment (x-axis). Each dot thus represents one participant, the participant with the number indicated at the x-axis. A negative coefficient implies that the participant tended to assign penultimate stress more often to quadrisyllabic than to trisyllabic words, while a positive coefficient indicates a reverse length effect. The larger the absolute value of the coefficient, the more word length affected the participant's responses. The difference between the participants with respect to their sensitivity to the phonological structure of the final syllable is less extreme: All participants assigned penultimate stress more often to words ending in open syllables.

The analysis with experimental word as random variable showed that also the words were affected differently by Word length (log-likelihood ratio = 147.19, $p < 0.001$) and by the Phonological structure of the final syllable (log-likelihood ratio = 89.35, $p < 0.001$). In contrast to all other 27 words, three words (those which have the forms badozevɑx, fobetozɑs, pevabekɔf in the four-syllable condition with closed final syllables) tended to receive penultimate stress more often in their trisyllabic than in their quadrisyllabic forms. With respect to the Phonological structure of the final syllable, all words received penultimate stress more often when their final syllable was open. Only the effect size differed for each experimental word.

On average, generalizing across participants, we found the hypothesized effect of Word length on the location of primary stress. Participants assigned penultimate stress more often to quadrisyllabic words than to trisyllabic words. We did not find an interaction of Word length by the Phonological structure of the final syllable: Word length was equally important for words ending in closed and in open syllables. Thus, although the words ending in open syllables did not show a length effect in the lexical analysis (Section 2.2), they did in the paper-and-pencil experiment. Recall that a relatively high percentage of the existing trisyllabic words of this type have initial primary stress (cf. Table 1), and we therefore could expect an effect of Word length also for these words.

3.3. *German*

We now turn to German. The analysis of the German monomorphemic words in Section 2.3 did not reveal any effect of the number of syllables in the word on the location of primary stress. This may be due to the low number of words in the data set, and we carried out a paper-and-pencil test with pseudowords also with native speakers of German. This experiment was identical to the Dutch experiment, except that the pseudowords conformed to the phonotactics of German, and that we spelled the words in the conventional spelling of German, rather than in the International Phonetic Alphabet, because we did not know whether all our participants would be sufficiently familiar with the International Phonetic Alphabet. We did not expect a difference between words ending in open and in closed syllables, as our lexical analysis classified these two types of words together for primary stress, and since also the literature provides no reasons to expect a difference between these two word types.

3.3.1. *Materials.* We created 40 quadrisyllabic words obeying the phonotactic restrictions of German. Their three initial syllables were open, while their final syllable was closed (e.g., *Zautaputasch*). From these 40 words, we created forty quadrisyllabic items ending in open syllables (*Zautaputa*), 40 trisyllabic items ending in closed syllables (*Taputasch*), and 40 trisyllabic items ending in open syllables (*Taputa*), along the lines described for the Dutch experiment. We created twelve experimental lists with these items, also following the procedure described for the Dutch experiment. The items were spelled in the conventional spelling of German (cf. the Appendix).

3.3.2. *Participants and procedure.* Forty-eight native speakers of German, recruited at the Universities of Cologne, Duisburg, and Kiel, participated in the experiment. The procedure was identical to that of the Dutch experiment.

3.3.3. *Results and discussion.* All participants indicated their preferred stress patterns for all experimental items in their lists. Table 10 lists the absolute numbers and percentages of responses with antepenultimate, penultimate, and final stress broken down for item type. We analyzed the penultimate stress responses versus the non-penultimate stress responses again by means of generalized linear mixed models, one with participant and one with experimental word as random variable. In both models, our independent variables were Word length (trisyllabic or quadrisyllabic) and the Phonological structure of the final syllable (open or closed). Both models showed a significant main effect for the Phonological structure of the final syllable (in the model with participant as random variable $F(1, 1869) = 14.31$, $p < 0.001$; and in the model with experimental word as random variable $F(1, 1877) = 15.29$, $p < 0.001$). Items ending in open syllables received more often penultimate stress than items ending in closed syllables. Word length was also significant, although marginally

Table 10. *The absolute numbers and percentages of German responses with antepenultimate, penultimate, and final stress for the quadrisyllabic and trisyllabic items with closed and open final syllables*

Item type		Location of primary stress					
Final syllable	Word length	Antepenult		Penult		Final	
closed	3 syllables	223	39.5%	200	35.5%	141	25.0%
closed	4 syllables	202	34.1%	278	47.0%	112	18.9%
open	3 syllables	163	26.8%	321	52.8%	124	20.4%
open	4 syllables	164	27.2%	333	55.2%	106	17.6%

in the model with participant as random variable (participant as random variable $F(1, 1869) = 3.60$, $p = 0.06$; experimental word as random variable $F(1, 1877) = 15.76$, $p < 0.01$). Participants assigned penultimate stress more often to quadrisyllabic than to trisyllabic words. Neither model showed an interaction of Word length by Phonological structure of the final syllable ($F(1, 1869) = 2.92$, $p > 0.1$; $F(1, 1877) = 1.58$, $p > 0.1$).

Participants differed in their sensitivity to the Phonological structure of the final syllable (Log likelihood ratio = 24.18, $p < 0.001$). Four participants tended to assign penultimate stress more often to words ending in closed than in open syllables. More importantly for our research question, participants also differed in their sensitivity to Word length (Log likelihood ratio = 27.73, $p < 0.001$). The random structure of the model showed that ten participants tended to assign penultimate stress more often to trisyllabic than to quadrisyllabic words. Figure 1b shows the coefficients for Word length for every participant in the experiment. A negative coefficient implies that the participant tended to assign penultimate stress more often to quadrisyllabic than to trisyllabic words and a positive coefficient the reverse.

The experimental words also differed in their sensitivity to the Phonological structure of the final syllable (Log likelihood ratio = 13.80, $p = 0.001$): Four items (those which have the forms *Bäwotapesch, Satodawett, Teifabatepp, Zaupotapitt* in the four-syllable condition with closed final syllables) tended to receive penultimate stress more often when their final syllable was closed. In addition, the items were also affected differently by Word length (Log likelihood ratio = 20.91, p < 0.001): Two items (*Fäbeetosass* and *Vauschateesock*) received penultimate stress more often when they were presented in their trisyllabic form.

We conclude that word length affected the location of primary stress in German pseudowords. As in Dutch, penultimate stress was more frequent among the quadrisyllabic than among the trisyllabic words. Also as in Dutch, speakers differed in their sensitivity to word length as well as to the phonological structure of the final syllable.

3.4. *English*

The analysis of the English trisyllabic and quadrisyllabic existing words described in Section 2.4 showed that word length affects the location of primary stress in words with heavy and VC penultimate syllables. This analysis was mainly based on nouns, and given the results from the paper-and-pencil experiments in Dutch and German, we may therefore expect an effect of word length in English pseudo nouns of this phonological structure. In order to see whether the effect of word length extends to pseudo verbs, we presented our pseudowords in the paper-and-pencil experiment as verbs.

We constructed words which consisted of heavy syllables followed by a final syllable that was heavy in half of the cases, and a VC-syllable in the other half. We did not have clear predictions for the effect of the phonological structure of the final syllable. On the one hand, all words with heavy penultimate syllables were grouped together in the CART classifications in Table 8, independently of the phonological structure of the final syllable. This suggests that there may be no effect of the phonological structure of the final syllable. On the other hand, according to the literature (see Section 2.4), the difference between final heavy syllables and final VC syllables is relevant for verbs, and the participants were told that the words were verbs.

3.4.1. *Materials.* We constructed thirty-five pseudo verbs, obeying the phonotactic constraints of English. The verbs started with three open heavy syllables with long vowels and ended in another heavy syllable that contained a long vowel (e.g., *potabovoo*) or ended in a short vowel and two consonants. From these quadrisyllabic verbs, we created trisyllabic items by cutting off the initial syllable (*tabovoo*). In addition, we created quadrisyllabic versions ending in syllables with a short vowel and one coda consonant (*potabovock*), and their trisyllabic counterparts (*tabovock*). All experimental items are listed in the appendix. They were presented to participants in the conventional orthography of English, since our participants were not familiar with the International Phonetic Alphabet. Hence, the weight of the penultimate syllable as perceived by the participant could not be completely controlled for.

We created three master lists, each containing one version of every experimental word, and eight or nine tokens of every item type. From each of these lists, we created three complementary lists, such that the complementary lists also contained eight or nine tokens of every item type, and together with the master list they contained all experimental items. The items in each list were randomized, but each experimental word occupied the same position in a master list and in its three corresponding lists. This procedure resulted in 12 experimental lists.

3.4.2. *Participants and procedure.* Forty-four native speakers of English, recruited at the University of Birmingham, participated in the experiment. They had not received any education in stress theory or the stress rules of English. The procedure was identical to the procedure of

Table 11. *The absolute numbers and percentages of English responses with antepenultimate, penultimate, and final stress for the quadrisyllabic and trisyllabic items with heavy and VC final syllables*

Item type		Location of primary stress					
Final syllable	Word length	Antepenult		Penult		Final	
heavy	3 syllables	186	48.3%	86	22.3%	113	29.4%
heavy	4 syllables	166	43.6%	127	33.3%	88	23.1%
VC	3 syllables	174	45.5%	108	28.3%	100	26.2%
VC	4 syllables	159	41.5%	137	35.8%	87	22.7%

the experiments in Dutch and German, except that the participants were told that the pseudowords were verbs.

3.4.3. *Results and discussion.* Two participants did not indicate their preferred stress pattern for one experimental item, while another participant refrained from giving responses for seven experimental items. Table 11 lists the absolute numbers and percentages of responses with antepenultimate, penultimate, and final stress broken for item type (quadrisyllabic or trisyllabic items with heavy or VC final syllables). We analyzed the penultimate stress and non-penultimate stress responses as a function of Word length (trisyllabic or quadrisyllabic) and the Phonological structure of the final syllable (heavy or VC), again using generalized linear mixed models. We observed a significant main effect for Word length, both in the model with participant as random variable ($F(1, 1485) = 6.49$, $p = 0.01$) and in the model with experimental word as random variable ($F(1, 1494) = 15.57$, $p < 0.001$). The phonological structure of the final syllable was significant in the model with participant as random variable ($F(1, 1485) = 3.98$, $p < 0.05$) and marginally significant in the model with experimental word as random variable ($F(1, 1494) = 3.24$, $p = 0.07$). Neither model showed an interaction between these two independent variables ($Fs < 1$). Quadrisyllabic words elicited penultimate stress more often than trisyllabic words, and words ending in VC syllables seem to elicit penultimate stress slightly more often than words ending in heavy syllables. This latter difference suggests that the participants indeed interpreted the pseudowords as verbs. Hence, we can conclude that the length effect observed in nouns in Section 2.4 extends to verbs.

The participants differed in their sensitivity to the length of the word (log-likelihood ratio = 57.35, $p < 0.001$). The model with participant as random effect assigned positive coefficients for Word length to 16 out of the 40 participants (see Figure 1c), which suggests that these participants preferred penultimate stress in trisyllabic instead of quadrisyllabic words. The model with experimental word as random effect showed no difference between the words in their sensitivity to Word length, or the Phonological structure of the final syllable.

We conclude that speakers of English tend to be sensitive to word length for the location of primary stress, and to assign penultimate stress more often to quadrisyllabic than to trisyllabic words. These findings are in line with the results of the lexical analysis presented in Section 2.4.

4. Variation

The lexical analyses in Section 2 and the paper-and-pencil experiments discussed in Section 3 show different types of variation. The lexical analyses show variation in the location of primary stress among words of the same number of syllables and ending in syllables of the same phonological structure. Thus, among the trisyllabic Dutch existing words ending in two open syllables, $(63/(63 + 74 + 24) * 100\% =)$ 39% have antepenultimate primary stress, $(74/(63 + 74 + 24) * 100\% =)$ 46% have penultimate stress, and $(24/(63 + 74 + 24) =)$ 15% have final stress (cf. Table 1). A possible explanation, as suggested by Baker and Smith (1976), van der Hulst (1984), Kager (1989), Don and Zonneveld (1988), Daelemans et al. (1994), Wiese (1996), and Féry (1998), among others, is that not only the phonological structures of the final syllables are relevant for the location of primary stress, but also the segments forming these syllables. For instance, the position of primary stress in Dutch words consisting of two open syllables is co-determined by the quality of the vowel in the final syllable. A tense [e] in this syllable favors final stress, while a tense [a] favors initial stress (van der Hulst 1984: 220).

The paper-and-pencil experiments showed that there is interspeaker variation. In all three paper-and-pencil experiments, we found that the majority of participants showed the expected effect of word length on the position of primary stress, but also that some participants showed no effect at all, or an effect in the opposite direction (see Figure 1). Possibly, the way in which we presented the words to the participants may have contributed to this interspeaker variation. The Dutch pseudowords were presented in the International Phonetic Alphabet, but possibly the Dutch participants were not equally familiar with this alphabet. The English words were presented in the conventional orthography of English, although we could not completely control with this spelling the way the participants would pronounce the words. The German pseudowords were spelled in the conventional spelling of German, and this spelling possibly indicated most robustly our intended pronunciation of the words. The interspeaker variation was also smallest for the Germans (see Figure 1). The Germans, however, also showed the smallest overall effect for word length, which is an alternative explanation for the relatively small interspeaker variation in this language. A second explanation for the interspeaker variation may be that the participants who did not show the length effect or showed the reversed length effect may not have known the quadrisyllabic words in their language that are necessary to establish the length effect. That is, they might have lacked the necessary exemplars. Third, these speakers may be less sensitive to rhythm, and, as a consequence, also to factors affecting the position of stress.

The paper-and-pencil experiments also revealed intraspeaker variation. The coefficients for Word length assigned by the models represent the probability that the responses of a participant are affected by word length. We found that all participants have small coefficients, which implies that they show variation with respect to the word length effect. This intraspeaker variation poses a challenge to phonological theory, which has been taken up, among others, by Antilla (2002).

Antilla (2002) discussed roughly three ways to incorporate intraspeaker language variation into Optimality Theory. First, in the tied violations account, variation is considered to be the consequence of output candidates that perform equally well on the constraint hierarchy (tied candidates). This account is somewhat problematic, however, as pointed out by Antilla, since tied candidates are highly unlikely, given the large inventory of universal constraints which probably distinguishes between any two candidates.

A second account of intraspeaker variation assumes that all speakers may have several grammars at their disposal which differ only slightly in the ranking of some phonological constraints (multiple grammars account). A given speaker would choose a given candidate more often if this candidate is optimal in more of the speaker's grammars. In our opinion, this approach is unable to account for the data from our paper-and-pencil experiments in an elegant way. Figure 1 clearly shows that the coefficient for Word Length may have any value between -3 (indicating a relatively strong effect of Word length in the predicted direction) to 1 (indicating a weak opposite effect). In order to account for the small differences between some speakers and the large differences between others, we have to assume within the multiple grammars account that all speakers have a large number of different grammars available to them, and that speakers may differ from each other in just one of their grammars or in many of their grammars (depending on their difference in the coefficient, or effect size, of Word length). Theoretically, this is possible, but the assumption that each speaker has the disposal of a high number of grammars is not very attractive.

Stochastic Optimality Theory (Boersma 1998; Boersma and Hayes 2001), a third account of variation, solves this problem. Within Stochastic Optimality Theory, constraints vary slightly in their exact position in the constraint hierarchy. Consequently, a constraint that is generally ranked higher than another constraint may be ranked lower than this constraint in a certain percentage of cases, resulting in variation in the output. The more two constraints differ in their average rank in the hierarchy, the smaller the probability that the relevance of these constraints is sometimes reversed. The stochastic part of Stochastic Optimality Theory thus elegantly accounts for intraspeaker variation. Interspeaker variation results from differences between speakers in the exact position of the constraints in the hierarchy.

Alternatively, the inter- and intraspeaker variation in the locations of primary and secondary stresses may be accounted for in data-oriented approaches, in which the choice for the stress pattern for a given (new) word is based on the stress patterns in phonologically, morphologically, and semantically similar words (e.g., Baker and Smith 1976; Daelemans et al. 1994). According to these models, if a (new) word has to be assigned a stress pattern, all words in the mental lexicon that are similar to this word, on phonological, morphological, or semantic grounds, are activated. The probability that a certain stress pattern is chosen is estimated by the relative frequency of that stress pattern in the similarity neighborhood of activated words. In these models, interspeaker variation would result from differences between the exemplars present in the speakers' lexicons, that is, from differences between speakers in which words they have lexically stored, and possibly also in the strength of the lexical representations. The intraspeaker variation may result from the temporary absence of some words from the exemplar set due to oblivion.

Data-oriented approaches can also easily account for the observation that words sharing certain segments tend to carry primary stress on the same syllable (e.g., Baker and Smith 1976; van der Hulst 1984; Don and Zonneveld 1988; Kager 1989; Daelemans et al. 1994; Wiese 1996; Féry 1998; Hammond 1999). The activation of a word may be stronger if the similarities with the target word are greater. Words which do not only share their length and the phonological structure of their final syllables with the target word, but also several segments, may thus have a greater effect on the position of primary stress in the target word. Incorporating "segmental" effects in Optimality Theory, in contrast, is less straightforward. It calls for constraints which state that words with certain segments have to have primary stress on a given syllable. Such constraints are theoretically possible, but render the accounts cumbersome.

5. General discussion

This study addressed the question as to whether the location of primary stress in Dutch, German, and English monomorphemic words is affected by the preceding syllables in the word. These languages show preferences for word-initial syllables with primary or secondary stress and for alternating patterns of stressed and unstressed syllables. We hypothesized that these preferences may result in an effect on the position of primary stress. We predicted that penultimate stress (in combination with initial stress) would be more frequent among quadrisyllabic than trisyllabic words.

We carried out statistical analyses of the monomorphemic words included in the CELEX lexical database, taking into account the phonological structure of the three final syllables, as well as word class for English. In the German data set, we found no effect of word length, possibly because of data sparseness. Both the Dutch and the English data set showed that penultimate stress is less frequent among trisyllabic than quadrisyllabic words. The effect is restricted to word types of which the trisyllabic words tend to have antepenultimate stress and in which the penultimate syllable may bear primary stress without inducing unwellformed feet.

In a series of paper-and-pencil experiments, we investigated whether the observed effects of word length in existing words reflect a property of the contemporary grammars of the languages. We presented pseudowords to native speakers of the three languages, asking them to indicate for every word to which syllable they would assign primary stress. The pseudowords were of the phonological types for which we could expect a length effect, partly on the basis of the lexical analyses. The length effect was statistically significant in all three experiments. The analyses, however, also revealed that not all participants showed the predicted effect. Some participants showed no effect, while others showed an effect that was opposite to what we expected. In addition, we attested intraspeaker variation: Speakers did not assign the same stress pattern to all words of a certain type. Finally, we attested systematic differences between words.

We explained the effect of word length as being driven by a preference for initial stressed syllables and alternating patterns of stressed and unstressed syllables (henceforth: “initial stress account”). An alternative explanation of the length effect is also possible, however (Don and Zonneveld 1988). In quadrisyllabic words, a perfect parsing of the syllables into wellformed feet implies two trochees, one starting at the initial syllable, and one starting at the penultimate syllable. In trisyllabic words, in contrast, such a footing is not optimal as it results in a stress clash. In other words, the effect of word length on the location of primary stress may just result from the preference for parsing as many syllables as possible into well-formed feet. This perfect parsing account is more economical than the initial stress account in that it does not need the assumption of a preference for word-initial stress. This assumption, however, is supported by several psycholinguistic studies showing preferences for word-initial stress in language production and perception (e.g., Baker and Smith 1976; Cutler and Norris 1988; Slootweg and Rietveld 1989).

Remark that according to both explanations, the position of primary stress depends on the positions of the secondary stresses. We thus conclude that our findings call into question the assumption that the location of primary stress is independent of the location of secondary stresses.

Interestingly, our findings bring Dutch, German, and English more in line with those languages, which, like Dutch, German, and English, have binary weight-sensitive feet and primary stress in the final three-syllable window, but in which footing is assumed to start from the left, instead of the right edge (e.g., Palestinian Arabic and Creek, e.g., Hayes 1995). Also in these languages, the footing of the syllables before the final three-syllable window affects the position of primary stress.

Received 20 February 2004
Revised version received
10 June 2005

Radboud University Nijmegen
Max Planck Institute for Psycholinguistics

Appendix

Experimental items for the Dutch paper-and-pencil experiment

	4 syll. closed	4 syll. open	3 syll. closed	3 syll. open
1.	ketozadɑp	ketozada	tozadɑp	tozada
2.	ɣedafokɛt	ɣedafoke	dafokɛt	dafoke
3.	fakodaɣɔp	fakodaɣo	kodaɣɔp	kodaɣo
4.	baɣotofɑx	baɣotofa	ɣotofɑx	ɣotofa
5.	bafopasɛk	bafopase	fopasɛk	fopase
6.	petaboxɔf	petaboxo	taboxɔf	taboxo
7.	bepotavɛx	bepotave	potavɛx	potave
8.	pefatodɛt	pefatode	fatodɛt	fatode
9.	kaxofasɛp	kaxofase	xofasɛp	xofase
10.	defapokɛt	defapoke	fapokɛt	fapoke

11.	fapobebɔx	fapobebo	pobebɔx	pobebo
12.	fobetozɑs	fobetoza	betozɑs	boteza
13.	pedakotɑf	pedakota	dakotɑf	dakota
14.	bekosafɑk	bekosafa	kosafɑk	kosafa
15.	pesoɣofɑp	pesoɣofa	soɣofɑp	soɣofa
16.	vekatobɔx	vekatobo	katobɔx	katobo
17.	katokasɑf	katokasa	tokasɑf	tokasa
18.	patosapɔs	patosapo	tosapɔs	tosapo
19.	taɣopekɛt	taɣopeke	ɣopekɛt	ɣopeke
20.	zapotepɛk	zapotepe	potepɛk	potepe
21.	bavotepɛs	bavotepe	votepɛs	votepe
22.	badozevɑx	badozeva	dozevɑx	dozeva
23.	potaɣovɔk	potaɣovo	taɣovɔk	taɣovo
24.	kotebetɑf	kotebeta	tebetɑf	tebeta
25.	kaɣopotɔt	kaɣopoto	ɣopotɔt	ɣopoto
26.	ɣebatokɔs	ɣebatoko	batokɔs	batoko
27.	pevabekɔf	pevabeko	vabekɔf	vabeko
28.	tepovadɛs	tepovade	povadɛs	povade
29.	depobatɑx	depobata	pobatɑx	pobata
30.	ɣofabodɛs	ɣofabode	fabodɛs	fabode

Experimental items for the German paper-and-pencil experiment

	4 syll. closed	4 syll. open	3 syll. closed	3 syll. open
1.	Tapeedabapp	Tapeedaba	Peedabapp	Peedaba
2.	Deibeekaeto	Deibeekaetoss	Beekaeto	Beekaetoss
3.	Kautosadapp	Kautosada	Tosadapp	Tosada
4.	Scädafokiff	Scädafokie	Dafokiff	Dafokie
5.	Fakeedatopp	Fakeedato	Keedatopp	Keedato
6.	Bapfatupasch	Bapfatupa	Pfatupasch	Pfatupa
7.	Bafopaweck	Bafopawee	Fopaweck	Fopawee
8.	Pietabäschoff	Pietabäscho	Tabäschoff	Tabäscho
9.	Giepeetageck	Giepeetagee	Peetageck	Peetagee
10.	Päfatodesch	Päfatodee	Fatodesch	Fatodee
11.	Kaschofeesipp	Kaschofeesie	Schofeesipp	Schofeesie
12.	Dupfapakett	Dupfapakee	Pfapakett	Pfapakee
13.	Keifateedosch	Keifateedo	Fateedosch	Fateedo
14.	Zautaputasch	Zautaputa	Taputasch	Taputa
15.	Pfapodeebüsch	Pfapodeebü	Podeebüsch	Podeebü
16.	Fäbeetosass	Fäbeetosa	Beetosass	Beetosa
17.	Satodawett	Satodawee	Todawett	Todawee
18.	Feideekutasch	Feideekuta	Deekutasch	Deekuta
19.	Gaufusafack	Gaufusafa	Fusafack	Fusafa
20.	Peisogeefapp	Peisogeefa	Sogeefapp	Sogeefa
21.	Bauteekaschoss	Bauteekascho	Teekaschoss	Teekascho
22.	Zakogapiff	Zakogapie	Kogapiff	Kogapie

23.	Vauschateesock	Vauschateeso	Schateesock	Schateeso
24.	Fabokasaff	Fabokasa	Bokasaff	Bokasa
25.	Fateesaposs	Fateesapo	Teesaposs	Teesapo
26.	Tageepukesch	Tageepukee	Geepukesch	Geepukee
27.	Zaupotapitt	Zaupotapie	Potapitt	Potapie
28.	Seepfakiewock	Seepfakiewo	Pfakiewock	Pfakiewo
29.	Bäwotapesch	Bäwotapee	Wotapesch	Wotapee
30.	Beudieguwasch	Beudieguwa	Dieguwasch	Dieguwa
31.	Scätapawott	Scätapawo	Tapawott	Tapawo
32.	Kautabotüff	Kautabotü	Tabotüff	Tabotü
33.	Sceupodafütt	Sceupodafü	Podafütt	Podafü
34.	Sceidasawoss	Sceidasawo	Dasawoss	Dasawo
35.	Zeewabuwatt	Zeewabuwa	Wabuwatt	Wabuwa
36.	Teifabatepp	Teifabatee	Fabatepp	Fabatee
37.	Feipowadesch	Feipowadee	Powadesch	Powadee
38.	Diepobatack	Diepobata	Pobatack	Pobata
39.	Beepfiedawott	Beepfiedawo	Pfiedawott	Pfiedawo
40.	Fateebadoss	Fateebado	Teebadoss	Teebado

Experimental items for the English paper-and-pencil experiment

	4 syll. heavy	4 syll. VC	3 syll. heavy	3 syll. VC
1.	tasodapeat	tasodapish	posapeat	posapish
2.	dibokatiece	dibokatass	bokatiece	bokatass
3.	kotasodont	kotasodoth	tasodont	tasodoth
4.	thodafokent	thodafokith	dafokent	dafokith
5.	facodathay	facodathip	codathay	codathip
6.	bafopasike	bafopasit	fopasike	fopasit
7.	potabogatch	potabogat	tabogatch	tabogat
8.	vusotavate	vusotavap	cotavate	cotavap
9.	picatogaine	picatogock	catogaine	catogock
10.	capofothobe	capofotholl	pofothobe	pofotholl
11.	dafaponont	dafaponom	faponont	faponom
12.	stobatadoy	stobatador	fabadoy	batador
13.	kotasofasp	kotasofal	tasofasp	tasofal
14.	fapobifown	fapobifon	pobifown	pobifon
15.	sitadothike	sitadothull	tadothike	tadothull
16.	fadakifow	fadakifock	dakifow	dakifock
17.	pasogofimp	pasogofid	sogofimp	sogofid
18.	slethokapace	slethokapoll	thokapace	thokapoll
19.	clabothosape	clabothosal	bothosape	bothosal
20.	vodapodench	vodapodesh	dapodench	dapodesh
21.	grathogasay	grathogasish	thogasay	thogasish
22.	patosadite	patosadar	tosadite	tosadar
23.	clapotifoy	clapotifick	potifoy	potifick
24.	pathokutaft	pathokutaff	thokutaft	thokutaff

25.	blathopatatch	blathopatash	thopatatch	thopatash
26.	tradovidotch	tradovidiff	dovidotch	dovidiff
27.	potabovoo	potabovock	tabovoo	tabovock
28.	chotasidorn	chotasitom	tasidorn	tasitom
29.	pladopitidge	pladipatil	dopitidge	dipatil
30.	prabosokipe	prabosokip	bosokipe	bosokip
31.	dovabatorp	dovabatoth	vabatorp	vabatoth
32.	thrapovadote	thrapovadosh	povadote	povadosh
33.	dufobatope	dufobatock	fobatope	fobatock
34.	buvodopox	buvodopod	codopox	codopod
35.	fobasodoat	fobasodit	basodoat	basodit

Notes

* Correspondence address: Mirjam Ernestus, MPI for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands. E-mail: mirjam.ernestus@mpi.nl.

1. We would like to thank Mirjam Broersma, Claudia Kuzla, and Martin Neef for helping us with the experimental items, and for running the experiments. Furthermore, we thank Harald Baayen, Geert Booij, Carlos Gussenhoven, and two anonymous reviewers for their comments.

References

Antilla, Arto (2002). Variation and phonological theory. In *The Handbook of Language Variation and Change*, John Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), 206–243. Oxford: Blackwell.

Baayen, Harald; Piepenbrock, Richard; and Gulikers, Leon (1995). *The CELEX Lexical Database (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Baker, Robert; and Smith, Philip (1976). A psycholinguistic study of English stress assignment rules. *Language and Speech* 19, 9–27.

Boersma, Paul (1998). *Functional Phonology*. The Hague: Holland Academic Graphics.

Boersma, Paul; and Hayes, Bruce (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32, 45–86.

Booij, Geert (1995). *The Phonology of Dutch*. Oxford: Clarendon Press.

Breiman, Leo; Friedman, Jerome; Olshen, Richard; and Stone, Charles (1984). *Classification and Regression Trees*. Belmont, CA: Wadsworth International Group.

Chomsky, Noam; and Halle, Morris (1968). *The Sound Pattern of English*. New York: Harper and Row.

Cutler, Anne; and Norris, Dennis (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance* 14, 113–121.

Daelemans, Walter; Gillis, Steven; and Durieux, Gert (1994). The acquisition of stress, a data-oriented approach. *Computational Linguistics* 20(3), 421–451.

Don, Jan; and Zonneveld, Wim (1988). VC-phonology. *Progress Report Institute of Phonetics Utrecht* 13, 8–32.

Féry, Caroline (1998). German word stress in Optimality Theory. *Journal of Comparative Germanic Linguistics* 2, 101–142.

Gussenhoven, Carlos (i.p.). Vowel duration, syllable quantity and stress in Dutch. In *The Nature of the Word: Studies in Honor of Paul Kiparsky*, Kristin Hanson and Sharon Inkelas (eds.). Cambridge, MA: MIT Press.

Halle, Morris; and Kenstowicz, Michael (1991). The free element condition and cyclic versus noncyclic stress. *Linguistic Inquiry* 22, 457–501.

Hammond, Mike (1999). *The Phonology of English: A Prosodic Optimality-Theoretic Approach*. Oxford: Oxford University Press.

Hayes, Bruce (1981). A metrical theory of stress rules. Unpublished doctoral dissertation, Yale University.

—(1982). Extrametricality and English stress. *Linguistic Inquiry* 13, 227–276.

—(1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.

van der Hulst, Harry (1984). *Syllable structure and stress in Dutch*. Dordrecht: Foris.

—(1985). Ambisyllabicity in Dutch. *Linguistics in the Netherlands 1985*, Hans Bennis and Frits Beukema (eds.), 57–66. Dordrecht: Foris.

van der Hulst, Harry; and Kooij, Jan (1992). Main stress and secondary stress: two modes of stress assignment. In *Phonologica. Proceedings of the 7th International Phonology Meeting*, Wolfgang Dressler, Martin Prinzhorn and John Rennison (eds.), 107–114. Turin: Rosenberg and Sellier.

Hung, Henrietta (1994). The rhythmic and prosodic organization of edge constituents. Unpublished doctoral dissertation, Brandeis University.

Janßen, Ulrike (2003). Wortakzent im Deutschen und Niederländischen, Experimentelle Untersuchungen zum deutschen und niederländischen Wortakzent. Unpublished doctoral dissertation, Heinrich Heine University, Düsseldorf.

Jensen, John (1993). *English Phonology*. Amsterdam: John Benjamins.

Jessen, Michael (1999). German. In *Word Prosodic Systems in the Languages of Europe*, Harry van der Hulst (ed.), 515–545. Berlin and New York: Mouton de Gruyter.

Kager, René (1989). *A Metrical Theory of Stress and Destressing in English and Dutch*. Dordrecht: Foris.

Martin, W. and Tops, G. A. J. (1997). *Van Dale Groot Woordenboek Engels-Nederlands*. Utrecht: Van Dale Lexicografie.

Neijt, Anneke and Zonneveld, Wim (1982). Metrische fonologie — de representatie van klem- toon in Nederlandse monomorfematische woorden. *De Nieuwe Taalgids* 75, 527–547.

Nouveau, Dominique (1994). *Language Acquisition, Metrical Theory, and Optimality: A Case Study of Dutch Word Stress*. Utrecht: LED.

Pater, Joe (2000). Non-uniformity in English secondary stress: the role of ranked and lexically specific constraints. *Phonology* 17, 237–274.

Prince, Alan (1983). Relating to the grid. *Linguistic Inquiry* 14, 19–100.

Prince, Alan and Smolensky, Paul (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell.

Schall, Robert (1991). Estimation in generalized linear models with random effects. *Biometrika* 78, 719–727.

Selkirk, Elisabeth (1984). *Phonology and Syntax*. Cambridge, MA: MIT Press.

Slootweg, Annemarie and Rietveld, Toni (1989). Naive subjects' assignments of stress ranks. *Phonetica* 46, 204–216.

Trommelen, Mieke (1991). Dutch word stress assignment: extrametricality and feet. In *The Berkeley Conference on Dutch Linguistics 1989: Issues and Controversies, Old and New*, Thomas Shannon and Johan Snapper (eds.), 157–172. Lanham, MD: University Press of America.

Trommelen, Mieke; and Zonneveld, Wim (1999a). English. In *Word Prosodic Systems in the Languages of Europe*, Harry van der Hulst (ed.), 478–491. Berlin and New York: Mouton de Gruyter.

—(1999b). Dutch. In *Word Prosodic Systems in the Languages of Europe*, Harry van der Hulst (ed.), 492–515. Berlin and New York: Mouton de Gruyter.

Venables, W. N.; and Ripley, B. D. (2002). *Modern Applied Statistics with S-Plus*, 4th ed. Berlin and New York: Springer.

Venneman, Theo (1990). Syllable structure and simple accent in Modern Standard German. In *Papers from the 26th Regional Meeting of the Chicago Linguistic Society: Volume 2. The Parasession on the Syllable in Phonetics and Phonology*, Karen Denton, Manuela Noske, and Michael Ziolkowski (eds.), 399–412. Chicago: Chicago Linguistic Society.

Wiese, Richard (1996). *The Phonology of German*. Oxford: Clarendon Press.

Zonneveld, Wim; and Nouveau, Dominique (2004). Child word stress competence: an experimental approach. In *Constraints in Phonological Acquisition*, René Kager, Joe Pater, and Wim Zonneveld (eds.), 369–408. Cambridge: Cambridge University Press.