Stress Patterns of German Cardinal Numbers

Sarah M. Creer* and **Maria Wolters**†

*Theoretical and Applied Linguistics, U of Edinburgh, 40 George Square, GB-EH8 9LL Edinburgh

† Rhetorical Systems, 4 Crichton's Close, GB-EH8 8DT Edinburgh
maria.wolters@rhetorical.com

ABSTRACT

The German cardinal numbers from 20 to 99 consist of two parts: a lexeme for the unit followed by a lexeme for the decade. Both parts may receive primary stress. Since numbers are very frequent in texts, it is important for a text-to-speech system to correctly predict the location of primary stress on a given number. Based on a production and a perception experiment, we suggest a simple, phonologically plausible algorithm for stress assignment: if a number is phrase-final, then primary stress falls on the decade, otherwise, it falls on the unit.

1 Motivation

Numbers occur frequently in the input of text-to-speech (TTS) systems. Since they tend to convey new information, it is important to pronounce them correctly. German numbers greater than 20 present a special challenge: they are compounds that can receive primary stress in several different places.

Our long-term aim is to devise a stress assignment algorithm for all numbers. In this paper, we begin with the base case, numbers between 20 and 99. Each number is a single orthographical word, which consists of the unit, followed by the conjunction *und* (and) and the decade. For multiples of ten, the leading unit is omitted.

When predicting stress for German numbers, we have to consider two aspects: the position of primary stress and the alternation between primary and secondary stress.

In single-lexeme words, stress can fall on the final syllable, the penultima, or the ante-penultima [3]. When two lexemes words are joined to form a compound, this pattern is reversed and primary stress moves to the beginning of the first lexeme [8]. But this is not the pattern that we see in the citation form of numbers: For these words, primary stress is on the second lexeme, the decade, and not on the first lexeme, the unit. In running speech, however, primary stress is frequently switched to the unit, in line with the preferred stress pattern for compounds. In this paper, we investigate the conditions under which this switch happens. The prime suspect is stress clash: If a cardinal is followed by a word where primary precedes secondary stress, primary stress should be moved from the decade to the unit.

In addition to primary stress, roughly every second syllable tends to receive at least secondary stress [3]. For cardinal numbers, this means that if the unit receives primary stress, then the decade should receive at least secondary stress, and vice versa.

The aim of this paper is to determine whether stress clash affects the position of primary and secondary stress for German cardinal numbers. The paper itself is structured as follows: In Section 2, we present the potential stress class environments that were investigated. Our experiments are discussed in Section 3. The results suggest that the main influence on the location of primary stress is not stress clash, but the position of the number in a phonological phrase. To conclude, in Section 4 we sketch a simple algorithm for assigning stress to cardinals.

2 Potential Environments for Stress Clash

Based on a review of the literature on stress clash, we identified three conditions under which we would expect primary stress to shift to the unit. The first condition is classic Eurythmy as discussed at length in Metrical Phonology (for a summary, see e.g. [5]):

C1 - Strict Eurythmy: Primary stress will move to the unit if the primary stress of the following word precedes its secondary stress.

However, in a sentence, some primary accents are more perceptually prominent than others. A clash with a strong primary accent might therefore lead to a stress shift on the preceding numeral, whereas a clash with a weaker primary accent might leave the numeral unaffected. Wagner [7] has found that in German, nouns, proper nouns, and numerals tend to be more perceptually prominent than adjectives and adverbs. This leads us to our second condition:

C2 - Eurythmy + POS: Primary stress will move to the unit if the primary stress of the following word precedes its secondary stress, and if that word is a noun, proper name or numeral.

Finally, Varga [6] has shown that in doubly accented Hungarian compounds, accent deletion rules are blocked when this rule would lead to a gap of four or more unstressed sylla-

bles. With our third condition, we tested whether the stress shift on German numerals might be sensitive to the size of the resulting gap between primary stresses.

C3 - Eurythmy + Syllable Count: Primary stress will move to the unit if the primary stress of the following word precedes its secondary stress, and if the unit is monosyllabic. If the unit is disyllabic, then the stressed syllable of the unit is separated from the stressed syllable of the decade by two unstressed beats.

3 Experiments

3.1 Design

Testing the three hypotheses against each other is difficult, because there is no case in which C2 and C3 would predict a stress shift, but C1 would not. Therefore, we designed four sentence patterns that cover all four possible combinations of the three conditions.

Pattern snn: Here, the cardinal is separated from the following word by a minor phrase boundary. The cardinal constitues a separate noun phrase with an empty head, and the following discourse unit is an elaboration of that NP. Only very strict eurythmy would predict a shift here; in English, the phonological phrase boundary would neutralise any potential stress clashes [2].

Pattern sns: Here, the cardinal is followed by another modifier, an adjective; semantically, the cardinal modifies the adjective-noun unit. In this case, only eurythmy and syllable-restricted eurythmy predict a shift, but not POS.

Pattern ssn: Here, the unit consists of the number *sieben* (7), the only cardinal smaller than ten that has two syllables. *Sieben* is stressed on the first syllable. Syntactically, the cardinal is a modifier of a full NP; it is immediately followed by the nominal head. Here, only eurythmy (C1) and the POS restriction (C2) would predict a shift, but not the syllable restriction (C3), because there are two syllables between the first syllable of "sieben" and the first syllable of the unit, and more than four syllables between the two primary stresses.

Pattern sss: Here, the cardinal is immediately followed by the nominal head it modifies. The primary stress of that noun is on the first or second syllable. This setup is summarised in Table 1.

| Pattern | Hypothesis predicts primary stress on | | |
	C1	C2	C3
snn	unit	decade	decade
sns	unit	decade	unit
ssn	unit	unit	decade
sss	unit	unit	unit
nnn	decade	decade	decade

Table 1: Summary of patterns used in the experiment

In order to determine whether a stress shift has taken place, we need to compare the realisation of the same cardinal under two conditions: no-shift (nnn) and potential shift (sss, sns, ssn, snn). As a consequence, each pattern consists of two parts. Both parts share the same cardinal. In the first part, the cardinal occurs in a context that does not trigger stress shift: It is followed by a word where secondary stress precedes primary stress. We call this control condition nnn. In the second part, the cardinal occurs in one of the four stress clash conditions, snn, sns, ssn, and sss. We can then determine whether stress shift has occurred by comparing the realization of the cardinal in the nnn condition to its realisation in the stress clash condition. If the unit has been made more prominent, and the decade less prominent, then we have found evidence that primary stress has shifted to the unit.

Once we have established the conditions in which stress shift occurs, we need to interpret our findings. If we find a shift only in conditions ssn and sss, but not in conditions sns and snn, then this is evidence for C2, the POS restriction. If there is a shift in sns and sss, but not in ssn and snn, this points to C3: the number of syllables needs to be taken into account. If we see a shift in sns, ssn, and sns, then the more general C1 holds. A shift in all four conditions would imply that eurythmy is even observed across minor phrase boundaries.

3.2 Data

Each instance of a pattern consists of two question-answer pairs. The question is always a plain Wh-question that cues intonational focus on the object NP of the answer. The answer is a simple subject-verb-object sentence, where the object NP contains the cardinal we are interested in. Both question-answer pairs share the same cardinal. Here is a sample instance:[1]

Pair 1 - stress on decade:
Q: Was haben Sie geSEHen?   A: Ich habe
Q: What have you seen?   A: I have
siebenundzwanzig   *De*monstraTIONen   gesehen.
twenty-seven   demonstrations   seen.

Pair 2 - potential stress on unit:
Q: Was haben Sie geSEHen?   A: Ich habe
Q: What have you seen?   A: I have
siebenundzwanzig AUSser*ir*dische geSEHen.
twenty-seven   extraterrestrials seen.

In the first pair of each instance, the cardinal is followed by a noun with primary stress on the third or fourth syllable, and secondary stress on the first or second syllable. This is the no-clash condition nnn. In the second pair, the cardinal appears in a condition where stress shift could occur: snn, sns, ssn, or sss.

For each of the four patterns, we designed three instances, yielding a total of 24 question-answer pairs (4 patterns × 3 instances × 2 question-answer pairs per instance = 24). These pairs were mixed with 12 distractor question-answer pairs and 12 question-answer pairs that were used to inves-

---

[1]primary stress marked with CAPS, secondary stress with *italics*; no stress marked on cardinal

tigate the effect of contrastive stress on the accentuation of cardinals.

In order to establish which stress pattern speakers use for the citation form of cardinals, and which stress pattern is used phrase-finally, we recorded four cardinals, *fünfundzwanzig* (25), *sechsundfünfzig* (56), *siebenundsiebzig* (77), and *fünfundsechzig* (65) in two different contexts: phrase-final as an NP with an empty head, and phrase medial in their citation form. The sentence focus was always on the cardinal. The sentence patterns are as follows:
Phrase-final:
Ich KENne siebenundsiebzig.
I know seventy-seven.
Citation form:
Das heißt siebenundsiebzig auf DEUtsch.
This is-called seventy-seven in German.

The resulting 8 utterances (2 contexts × 4 numbers = 8 utterances) were interspersed with 8 distractor utterances, four for the phrase-final and four for the phrase-medial pattern.

Since one of the acoustic measures analysed was duration, the syllable nuclei of both unit and decade had to be easy to segment. Therefore, we chose lexemes where the stressed vowel was a monophthong preceded by an obstruent. For the units, we used *fünf* (5), *sechs* (6), and *sieben* (7); for the decades, *zwanzig* (20), *fünfzig* (50), *sechzig* (60), and *siebzig* (70).

3.3 Method

It is difficult to reduce perceived prominence to acoustic measures. Therefore, we opted for a two-pronged approach: In a production experiment, we recorded subjects reading the sentence patterns and analysed that data for acoustic correlates of stress. A perception experiment was then conducted to validate the results of the production study.

3.3.1 Production Experiment: The subjects of our production experiment were eleven native speakers of German, none of whom reported a speech disorder. The subjects were between 23 and 34 years old and came from different parts of Germany. None of the subjects guessed the purpose of the experiment.

Each of the 16 sentences and 48 question-answer pairs was printed on a separate A4 sheet of paper in large typeface. Subjects were instructed to read the utterance(s) first silently, making sure they understood them, then aloud. The sentences were first randomised, then checked to make sure that cardinals were not followed by cardinals with the same decade or unit.

All acoustic measurements were restricted to the nucleus of the stressed syllable of the unit, and the nucleus of the stressed syllable of the decade.[2] We decided to focus on the nuclei because the nucleus is the syllable constituent that shows the strongest effects of stress. The boundaries of

each nucleus were determined by hand by one of the authors. Then, using the ESPS toolkit[3], the following measures were extracted: duration, F0 range (defined as maximum F0-minimum F0), mean F0, maximum F0, and intensity. In order to smooth out variation between speakers, intensity was measured in dB and z-scores were computed for F0. Increases in all these measures, especially duration and F0, strongly correlate with stress [4].

Eleven utterances had to be discarded because of audio problems, misreadings, wrong phrasing, and hesitations or repairs on or shortly after the number.

3.3.2 Perception Experiment: To check whether stress shifts detected by the acoustic measures actually correspond to perceived changes in stress patterns, we asked three phonetically trained German native speakers to mark the perceived prominence of unit and decade for a subset of the data collected in the production experiment. As marking the data from all speakers would have been too time-consuming, we restricted ourselves to three female speakers, one with a low F0 range, one with a medium F0 range, and one with a large F0 range.

The listeners were asked to mark the relative prominence of the unit and the decade of each number by drawing a line in two boxes: the longer the line, the higher the prominence. For each cardinal, the listeners heard the entire sentence. The sentences were presented in three blocks, one per speaker; within each block, the sequence of the sentences was randomised. The first six sentences were repeated again at the end of the block, and the first six judgements were discarded.

The lines were then converted into an interval scaled variable with discrete values by dividing all boxes into 7 equal segments, numbered from 0 to 6, where 6 corresponds to the highest, 0 to the lowest prominence.

3.4 Results

3.4.1 Production Experiment: In order to establish whether the stress pattern changes in any of the four potential stress-clash conditions, we conducted paired t-tests for all three instances of each of the four patterns.[4] In the t-tests, we compared the acoustic realisation of the unit and of the decade in the no-clash condition (nnn) with their realisation in the clash-condition snn, sns, ssn, and sss. For each test, we had 27 to 32 data points (maximum: 3 instances per pattern × 11 speakers = 33 data points). The significance level for all t-tests was 0.05.

Overall, we found no significant differences between any of the clash conditions and the corresponding no-clash utterances. There were no signs of any increase in the values of acoustic correlates of stress on the unit, and no signs of any decrease in those values on the decade. Figure 1 shows the difference between the duration of the unit nuclei for all four

---

[2]This might distort the result of F0 measures, since F0 peaks are aligned late in German [1].

[3]available from http://www.speech.kth.se/software/
[4]Inspection of normal quantile-quantile plots for the normalised variables and duration shows that they have a roughly normal distribution.
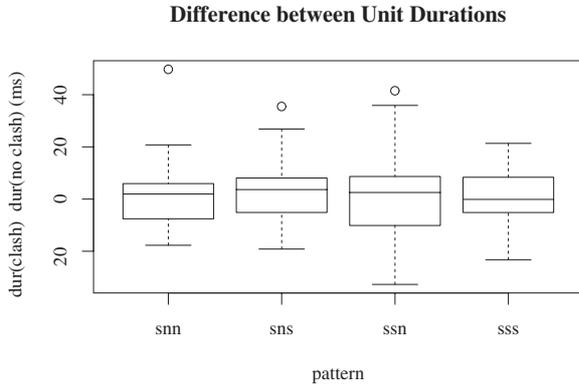
Difference between Unit Durations

Figure 1: Difference between duration of unit nuclei

conditions. The mean difference is close to 10 ms, almost imperceptible. This indicates that it is not stress clash that affects the stress pattern of German cardinals.

3.4.2 Perception Experiment: We analysed the difference between the prominence of the decade and the prominence of the unit, as judged by our three listeners. Positive values indicate that the unit was judged to be more prominent than the decade, negative values that the decade was more prominent than the unit, and zero shows that both were perceived as equally accented.

The listeners' judgements are highly correlated (Spearman's Rho, $p < 0.001$), although there are clear preferences: listener B tends to perceive the decade as more prominent and listener C the unit, while for listener I, both tend to be equally prominent. Pooling the results of all three listeners, we find that listeners tend to perceive stress on the decade if the number is phrase-final (pattern snn and phrase-final utterances) or used in its citation form, and stress on the unit if the number is phrase-medial (patterns ssn, sss, sns).[5]

more prominent	phrase medial	phrase final	citation
decade	13	36	13
neither	62	29	7
unit	89	13	4

Table 2: Perception of stress depending on position in the phrase

4 Conclusion

In phrase-medial position, cardinals behave much like normal compounds, with primary stress on the first component, the unit. This seems to be the default stress pattern. Primary stress does not even shift to the decade when the first stress of the following word is secondary. Instead, stress might be used to demarcate the boundaries of phonological phrases: phrase-finally, primary stress shifts to the decade. This is

exactly the behaviour we see in the citation form, where the cardinal forms a phrase of its own. When in doubt, the cardinal should receive two primary stresses, a pattern that was very frequent in the listeners' judgements.

This observation leads to an elegant stress assignment algorithm for the cardinal numbers between 20 and 99: If the cardinal occurs at the end of a phonological phrase, primary stress is on the decade, otherwise, it is on the unit.

It remains to be seen how this algorithm can be extended to numbers between 100 and 10000, with up to four distinct components. It would also be interesting to see whether other German words with two potential landing sites for primary stress behave in a similar way.

REFERENCES

[1] Michaela Atterer and D. Robert Ladd. On the phonetics and phonology of "segmental anchoring" of F0: evidence from German. *Journal of Phonetics*, 2002.

[2] Carlos Gussenhoven. The English Rhythm Rule as an Accent Deletion Rule. *Phonology*, 8:1–35, 1991.

[3] Michael Jessen. Word-stress in West Germanic and North-Germanic Languages: German. In Harry van der Hulst, editor, *Word Prosodic Systems in the Languages of Europe*, pages 515–545. Mouton de Gruyter, Berlin, 1999.

[4] Michael Jessen, Krzysztof Marasek, K. Schneider, and K. Clahssen. Acoustic correlates of word stress and the tense/lax opposition in the vowel system of German. In *Proc. 13th ICPhS*, volume 4, pages 428–431, Stockholm, 1995.

[5] Michael Kenstowicz. *Phonology in Generative Grammar*. MIT Press, Cambridge, MA, 1995.

[6] L. Varga. Rhythmical Variation in Hungarian. *Phonology*, 15:227–266, 1998.

[7] Petra Wagner. Rhythmical Alternations in German Read Speech. In *Proceedings of Prosody 2000, Krakow*, 2000.

[8] Richard Wiese. *The Phonology of German*. Cambridge University Press, Cambridge, UK, 1996.

---

[5]No statistical analysis was performed because this is a post-hoc result which will need to be verified in a new experiment.