Production of word stress in German: Children and adults

Katrin Schneider & Bernd Möbius

Institute of Natural Language Processing
University of Stuttgart, Germany
{katrin.schneider;bernd.moebius}@ims.uni-stuttgart.de

Abstract

This study investigates the acoustic correlates of contrastive word stress in bisyllabic and trisyllabic German words, produced by children and their parents. Results of the acoustic analysis of speech data are reported that were collected from three children aged 2;3 to 6;1 and their mothers during a period of two years, as a part of a more comprehensive study on the acquisition of word stress in German. Whereas recent findings suggest that infants show an early perceptual preference for rhythmic patterns of their native language, contrastive stress is supposed to be acquired relatively late. The results of the study presented here suggest that German children between 2 and 6 years of age are able to produce contrastive word stress but differ in their choice and usage of the parameters that mark stress. We found that, for German, vowel duration is the most reliable correlate of word stress in the utterances produced by all three children as well as their mothers. Adult-like usage of fundamental frequency, intensity, and several voice quality parameters appears to be acquired later than that of duration; this observation may be confounded by the finding that these parameters appear to be used less consistently than duration to mark stress even by the mothers.

1. Introduction

Perception studies show that prelinguistic infants are sensitive to the suprasegmental structure of their native language [8]. By 9 months of age, infants show a selective preference for stress patterns of their native language, and before the onset of meaningful speech, infants are sensitive to the frequency of the sound patterns in the ambient language [6]. Infants have to be sensitive to their native language's sound organization, its phonotactic structure, and its word structure, in order to segment words in fluent speech [14]. Infants rely on the prosodic characteristics of their native language to infer its syntactic properties. Variability makes it difficult to build up phonemic representations of the sounds of the native language. The acoustic realization of phonemes varies, depending on the speaker, speech rate or phonetic context, but listeners are able to assign different realizations of a given sound to the same category, a concept which is often referred to as phonemic constancy.

Many infant speech perception studies suggest that infants store specific information about voices and the words which they are exposed to. By using the input from their parents, babies establish expectations about the typical patterns of prosodic variations of utterances in the ambient language. They obtain sufficient exemplar variance to develop language specific generalizations about the distribution of values in the auditory dimensions of the phonetic space [4]. Each familiar word form is encoded in terms of episodic memory traces of fine-grained parametric representations of the auditory and articulatory patterns. The relationship of phonetic cues to the hierarchical prosodic and intonational representation must be learned for each individual language [9]. Proponents of Exemplar Theory assume that speech categories are stored as clouds of perceived exemplars in long-term memory. Speech perception and production are closely connected: during perception, the phonetic parameters of the input determine its placement in the phonetic space; during production, a set of exemplars of the respective category is randomly selected, and the median of this subset is calculated and produced with a random error [10].

To develop an exemplar-based model of the acquisition of contrastive stress in German we analyzed the effects of stress on its main acoustic correlates, i.e. on fundamental frequency (F_0), duration, and vowel quality. Contrastive stress is supposed to be acquired rather late in language development [2, 11]. We aim to investigate when children are able to produce contrastive stress and how they realize it. Our hypothesis is that children first adopt the features that their parents use to indicate stress, but we suppose that during later developmental stages they will be increasingly influenced by the ways other members of the local and regional linguistic community use to mark word stress.

2. Method

2.1. Participants

The data reported here are part of a larger investigation of children's acquisition of stress. We present the data collected during a period of two years and produced by three children, two brothers (aged 2;5 to 4;2 and 4;1 to 6;1) and one girl (aged 2;3 to 3;2). All three children live in monolingual German-speaking families. They had no unusual prenatal, sensory or developmental concerns, or hearing problems. Their parents were recorded too.

2.2. Stimuli

We adopted the TAKI task design proposed by Allen [1] and created five object pairs consisting of animal toys with bisyllabic or trisyllabic names. The names consist of consonant-vowel (CV) syllables and contain only speech sounds that are acquired as the first ones by German children: the vowels /a/, /i/, and /o/, and the consonants /b/, /d/, /m/, and /n/. The CV syllables are phonetically similar to reduplicated babbling and to a child's first words. The names within a pair of animals differ only in the position of word stress, e.g. /'bimo/ (a brown bear) vs. /bi'mo/ (a polar bear). In this contrastive stress condition, stressed and unstressed vowels can be compared in identical segmental contexts. The different stress positions in the target words were assigned such that they follow typical stress patterns of German words. In the CELEX database [3], for instance, the majority of German bisyllabic or trisyllabic words is

stressed on the first syllable, followed by words having stress on the second syllable. Stress on the third syllable is also possible, but this pattern is much less frequent. We analyzed the production of stress contrasts between the first and the second and the first and the third syllable of a word. The names of the animal toys are listed in Table 1.

Animal toy	Name
brown bear	Bimo ['bimo]
polar bear	Bimo [bi'mo]
big zebra	Nami ['nami]
small zebra	Nami [na'mi]
otter	Doba ['doba]
badger	Doba [do'ba]
big tiger	Midano ['midano]
small tiger	Midano [mida'no]
flying eagle	Badoni ['badoni]
standing eagle	Badoni [bado'ni]

Table 1: *Animal toys and their names with contrastive stress used in the TAKI task.*

2.3. Acoustic measurements

Each recording session was annotated on the segmental, syllabic, and word levels by two trained labelers. The absolute duration, F_0 contour, the first two formants and root mean square (RMS) intensity of each stressed and unstressed target vowel were measured. Voice quality parameters were estimated from their acoustic correlates based on amplitude and frequency measurements at harmonic spectral peaks [13]. The following voice quality parameters were extracted: open quotient (OQ), glottal opening (GO), skewness (SK), rate of closure (RC), completeness of closure (CC), and amplitude of voicing (AV). All measurements were made at five temporal positions within each target vowel (at 10%, 30%, 50%, 70%, and 90% of the total vowel duration) to capture parameter changes over the duration of the speech sound.

3. Results

For statistical analysis the SPSS program version 12.0 was used. A univariate generalized linear model (GLM) analysis was carried out for each parameter, followed by posthoc tests (Tukey-HSD and Waller-Duncan).

3.1. Parameters known to be independent of word stress

Results of previous studies suggest that adult speakers of German and Dutch do not rely on F_0 to mark word stress and that F_0 is vowel-specific [7, 12]. Our results generally support the vowel-specific characteristics of F_0 but, in addition, our data also show a dependency of F_0 on word stress, both for the children and their mothers. This effect is inconsistent as it differs between speakers and vowel phonemes. Furthermore, for all three children we found that F_0 significantly decreases with increasing age of the children.

The open quotient (OQ) is considered to be a correlate of sentence accent and is possibly caused by an additional modulation of muscular tension [12]. We found that OQ is vowel-specific and its dependency on stress differs between vowel phonemes and speakers. Both the children and their mothers show significantly higher OQ values for /a/ and /o/ than for /i/. This might support the assumption that OQ depends on muscle tension because back vowels tend to be produced with more muscular tension in the larynx than front vowels.

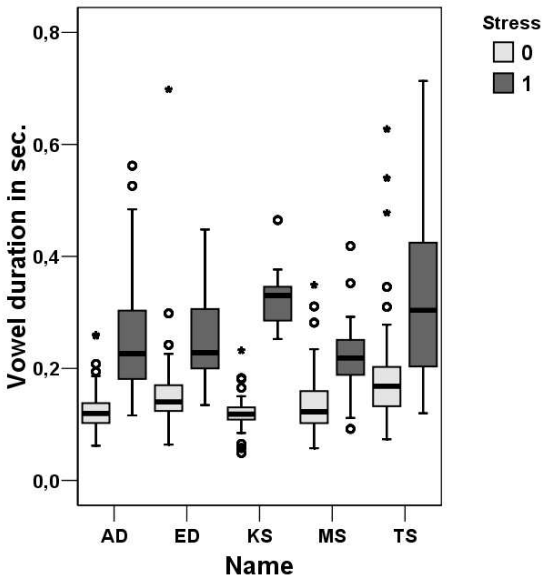

Figure 1: *Vowel duration for unstressed and stressed vowels produced by the mothers (AD and KS) and their children.*

3.2. Parameters depending on word stress

3.2.1. Vowel duration

Vowel duration has been shown to be the most reliable cue for word stress in German adults [7]. In our data, the duration of vowels differs strongly between vowel phonemes. We therefore transformed the absolute duration values of all vowel tokens to z-scores to normalize the durations across vowel phonemes.

For both the children and their mothers, stressed vowels are significantly longer than unstressed vowels (Figure 1). This effect is observed for all vowels pooled as well as for each vowel phoneme.

We also found evidence of lengthening of word-final syllables, but the strength of this final lengthening effect differs between speakers.

Our results show strong evidence for the fact that vowel duration is a reliable correlate of word stress in German children and adults. We also found that the total duration of vowels decreases with increasing age of the children.

3.2.2. Root mean square intensity

The girl's RMS intensity values for /i/ and /o/ vowels are significantly higher in stressed than in unstressed vowels ($p < 0.001$), but for /a/ vowels the opposite is observed. The RMS intensity values for the older boy show only a marginal significance ($p = 0.091$) for stress, with higher values for stressed than for unstressed /i/ and /o/ vowels. The younger boy's results show a significant dependency on stress for all vowels ($p < 0.01$), with unstressed vowels having lower intensity values than stressed vowels. Therefore, we suppose that the younger boy uses RMS intensity as a correlate of word stress throughout, whereas the older boy is inconsistent in his usage of intensity across vowels. It appears that the girl, the youngest child in this study, has

not yet detected the use of intensity for marking word stress in German. In general, she shows significantly higher RMS values than the two boys, which might be taken as a hint that she is not yet in the position to control this parameter for linguistic purposes. Furthermore, we found for all three children that RMS intensity depends on the age of the child. Overall, the two boys decrease their intensity values with increasing age, whereas the girl still appears to increase intensity with increasing age.

As for the adults, we found that the use of intensity is vowel-specific, yet inconsistent. Both speakers produce stressed /o/ vowels with higher RMS values than unstressed ones, but stressed /a/ is produced with higher intensity than unstressed /a/ only by one speaker and stressed /i/ is produced with higher intensity than unstressed /i/ only by the other speaker.

3.2.3. Glottal opening and completeness of closure

Glottal opening (GO) has been shown to be a correlate of stress in German adults [7]. Stressed vowels tend to have lower GO values than unstressed ones. These results are supported by our findings for the mothers' speech data. However, only the younger boy shows the adult-like significant patterns, whereas the other two children do this only for the /o/ vowels.

Sluijter found that stressed vowels show higher completeness of closure (CC) values, suggesting that they have a greater glottal leakage than unstressed vowels [12]. In our data we found the opposite tendencies for both adult speakers and all vowels. The children showed this opposite tendency for their /i/ and /o/ vowels.

3.2.4. Skewness and Rate of Closure

Spectral tilt has been shown to be a correlate of word stress in German adults [5]. Unstressed vowels tend to have a stronger spectral tilt than stressed ones. The same tendency was found for the related voice quality parameters, skewness (SK) and rate of closure (RC), even though these tended to be less consistent than spectral tilt.

The results of our study support the findings of Claßen et al. for the adult productions of the voice quality parameters SK and RC. Among the children, the younger boy shows the tendency to use SK in an adult-like manner, but this was significant only for two of the three vowel phonemes (/a/ and /o/). The girl's usage of SK appears to be the opposite of the adult-like use, while the older boy shows no clear tendency.

All three children use the RC parameter in an adult-like way for the vowel /o/, but they do not show a homogeneous tendency for the other two vowel phonemes. The girl produces the other vowels in opposition to the adult-like use. The younger boy produces the /i/ vowels with higher RC values for stressed vowels, while he uses the RC parameter adult-like for the /a/ vowels. The older boy shows a reversed use of the RC parameter for the /a/ and /i/ vowels compared to his younger brother.

3.2.5. Formants

Our results support the well-known finding that the first two formants are vowel-specific. For the adults in our study, F1 and F2 tend to depend on word stress, but with no homogeneous patterns across vowel phonemes and speakers.

For the older boy, all vowels have higher F1 and lower F2 values in stressed positions compared to unstressed positions. The other boy always produces higher formant values for unstressed /a/ vowels and always lower formant values for unstressed /o/ vowels compared to their stressed counterparts, but he varies in the formant use for the vowel /i/. The girl shows a constant pattern only for the vowel /a/, with higher formant values in stressed positions.

4. Discussion

The children's speech production data analyzed in this study provide clear evidence for the fact that children at the age of 2;6 are able to produce contrastive stress. However, the implementation and usage of the parameters that contribute to marking stress appear to be speaker-specific and somewhat inconsistent. Based on the hypothesis that perception leads production, we compared the acoustic correlates of word stress produced by these children to those produced by their mothers.

Our results confirm the findings reported in previous studies on adults' speech that the use of F_0 and the open quotient parameter (OQ) is vowel-specific. However, for both adults and children, we observed that F_0 as well as OQ can serve as correlates of word stress in German. There are several possible explanations. First, the target words in our study were not spoken in isolation or in otherwise controlled contexts, but produced in spontaneous mother-child interactions. Thus, sentence intonation and even discourse or dialog structure may have an effect on the realization of the target words that at present cannot be accounted for. Second, the mothers used child-directed speech, which is characterized, for instance, by a higher F_0 register and a larger F_0 range. A larger F_0 range may enhance the tonal difference between stressed and unstressed vowels in the target words. Finally, it is conceivable that child-directed speech is also characterized by voice quality settings that differ from those found in adult-directed speech.

Vowel duration is evidently the most reliable cue for word stress in German, in adults and even in children at the age of 2;6. All the subjects in our study used vowel duration to differentiate between stressed and unstressed vowels. In the absence of strong reduction effects as a function of stress in full vowels, vowel duration is one of the most salient cues for syllabic stress in German and may thus be implemented and used consistently even by young children.

Previous studies reported evidence that intensity is a correlate of word stress in German adults, with stressed vowels having higher RMS intensity values than unstressed ones. Our data do not completely support these findings for adults. The results for the two adult speakers in our study show RMS intensity values with the supposed dependency only for the /o/ vowels and one other vowel quality for each speaker. It is possible that child-directed speech also affects the intensity parameter. The children differed in the use of intensity to mark stress, too. Only the younger boy uses this parameter in the adult-like way described in previous studies. The girl, the youngest child in this study, may not yet have developed a consistent usage of vowel intensity. The older boy uses this parameter in the same way as his mother does.

With respect to the voice quality parameters, the results for the two adults in our study support previous findings that the glottal opening parameter (GO) depends on word stress, with unstressed vowels having higher GO values than stressed ones. In the children's data we found this pattern in all vowels produced by the younger boy but only in the /o/ vowels produced by the other children. Subsequent recordings will have to be analyzed to investigate whether these children will eventually acquire GO as a consistent stress cue.

The completeness of closure (CC) values extracted from our data suggest that unstressed vowels have a greater band-

width of the first formant than stressed vowels. This is in contrast to previous findings for German and Dutch adults [5, 12]. Sluijter explained her results by assuming that the subglottal pressure built up for stressed vowels also results in a greater glottal leakage. Our findings might be interpreted in two ways: either the production of stressed and unstressed vowels in German differs from that in Dutch with respect to the implementation of the CC voice quality parameter, or the differences between our results and former studies of adult language are due to the fact that we analyze child-directed speech produced by adults.

The parameters skewness (SK) and rate of closure (RC) describe the vibration behavior of the vocal folds. The closure of the vocal folds has been found to be more abrupt (SK) and faster (RC) in stressed than in unstressed vowels, and this difference is so consistent that these parameters are regarded as reliable cues for stress in German [5]. Our adult data support this conclusion: SK and RC show significantly higher values for unstressed than for stressed vowels. The results based on the children's speech data show that the children exploit these parameters which their parents present so consistently to them and have started to implement the SK and RC parameters according to the adult-like pattern.

5. Conclusion

The child data presented in this study suggest that children acquiring German as their native language use different strategies to produce word stress. The most reliable correlate of word stress in German children between the age of 2 and 6 years as well as for adults is vowel duration. Other stress cues reported for adults, including RMS intensity, skewness, rate of closure, glottal opening, and completeness of closure, already show tendencies of adult-like usage patterns, but these parameters are not yet stable. Differences between the results of previous studies of adult language and the results of the adult speech analyzed here may be partly due to the characteristics of child-directed speech used by the mothers while talking to their children. This special kind of speech may not only differ from adult-directed speech in fundamental frequency register and range, as reported in previous studies analyzing child-directed speech, but also in intensity and voice quality parameters. Experiments addressing the differences between adult-directed and child-directed speech are work in progress.

6. Acknowledgments

We thank Bernd Amann, Nicolai Hezinger, Stefan Kombrink and Manuel Kountz for labeling the speech data and programming the analysis tools, as well as the children and their parents for participating in the recordings. This study was carried out as part of the project An exemplar-theoretic model of the acquisition of acoustic correlates of stress funded by the German Research Foundation (DFG Grant MO 597/1).

7. References

[1] Allen, G.D., 1981. Development of prosodic phonology in children's speech: Further evidence from the TAKI task, In *Phonologica 1980*, Innsbrucker Beiträge zur Sprachwissenschaft 36, W.U. Dressler, O.E. Pfeiffer, J.R. Rennison (eds.), 9-14.

[2] Altmann, H., Kabak, B., 2000. The use of prosodic information for disambiguation by German children: An experimental investigation, Ms., University of Delaware.

[3] Baayen, H., Piepenbrock, R., Gulikers, L., 1995. The CELEX Lexical Database, Release 2, CD-ROM, Centre for Lexical Information, Max Planck Institute for Psycholinguistics, Nijmegen; Linguistic Data Consortium, University of Pennsylvania.

[4] Beckman. M.E., 2003. Input representations (Inside the mind and out), In *WCCFL 22 Proceedings*, G. Garding, M. Tsujimura (eds.), Somerville, MA: Cascadilla Press, 70-94.

[5] Claßen, K., Dogil,G., Jessen, M., Marasek, K., Wokurek, W., 1998. Stimmqualität und Wortbetonung im Deutschen, In *Linguistische Berichte* 174, Westdeutscher Verlag, 202-245.

[6] Curtin, S., Mintz, T.H., Christiansen, M.H., 2005. Stress changes the representational landscape: evidence from word segemtation, *Cognition*, 96, 233-262.

[7] Jessen, M., Marasek, K., Schneider, K., Claßen, K., 1995. Acoustic correlates of word stress and the tense/lax opposition in the vowel system of German, In *Proc. 13th ICPhS (Stockholm)* (4), 428-431.

[8] Jusczyk, P.W., Cutler, A., Redanz, N., 1993. Infants' preference for the predominant stress patterns of English words, *Child Development* 64, 675-687.

[9] Pierrehumbert, J.B., 2003. Probabilistic Phonology: Discrimination and Robustness, In *Probabilistic Linguistics*, R. Bod, J. Hay, S. Jannedy (eds.), MIT Press, Cambridge, MA.

[10] Pierrehumbert, J.B., 2003. Phonetic diversity, statistical learning and acquisition of phonology, *Language and Speech* 46 (2-3), 115-154.

[11] Pollock, K.E., Brammer, D.M., Hageman, C.F., 1993. An acoustic analysis of young children's productions of word stress, *J. of Phonetics* 21, 183-203.

[12] Sluijter, A.M.C., 1995. Phonetic correlates of stress and accent, PhD thesis, University of Leiden.

[13] Wokurek, W., Pützer, M., 2003. Automated corpus based spectral measurement of voice quality parameters, *Proc. 15th ICPhS (Barcelona)*, 2173-2176.

[14] Zamuner, T.S., Gerken, L., Hammond, M., 2004. Phonotactic probalities in young children's speech production, *J. Child Language* 31, 515-536.