

# Perception of Phrase Boundaries and Prominent Syllables in German

*Hansjörg Mixdorff, Angelika Hönemann, Hongwei Ding*

## Abstract

This study concerns the perception of boundaries and accented syllables by native German subjects as compared to Chinese non-speakers and learners of the language at different proficiency levels. To this effect six-syllable sequences excised from a context of three polysyllabic words of German were presented to participants who had to select the syllables they perceived as accented, as well as the locations of word boundaries. Results show that Germany native subjects perform well at the word boundary task, but mark correctly less than two thirds of accent syllables. Chinese non-learners still detect 70% of word boundaries and 46% of accent syllables. Learners of German - with the exception of a small group of advanced students after year 5 who performed even better than the native subjects - show growing improvement with respect to word boundaries, but mark even fewer accent syllables than the non-learners. Correlation analysis of factors for syllable and boundary selection performed for non-learners and German subjects - as expected - shows considerably different behaviors. Whereas the boundary location does not influence the Germans' decision on the accent location, Chinese non-learners show a preference to mark an accent when the syllable is followed by a word boundary.

## 1 Introduction

It is a well-known fact that non-speakers and native speakers of a language perceive and process stimuli from that language quite differently. When a person studies a foreign language (L2) (s)he makes a transition between the two states as (s)he acquires a growing competence as to L2 linguistic structures and vocabulary. An important competence for communication is the ability to process and chunk the speech stream into meaningful units. The segmentation of an utterance into words heavily depends on prosodic cues (see, for instance, Christophe et al., 2003), such as F0 and duration - features which might be employed quite differently in the L2 - as well as the growing L2 lexicon on the part of the learner.

Lexical stress is a property of each poly-syllabic word and in German - in contrast to certain other languages - its location in the word is rather flexible, i.e. not predefined by default (Kohler 1977). Hence three-syllable words, for instance, can exhibit lexical stress either on the first, second or third syllable. There exists a small group of words which are segmentally identical, but differ as to the lexical stress location (set in bold face): compare, for instance, 'um-**fah**-ren' (*to go around*) vs. '**um**-fah-ren' (*to run over*). In the context of an utterance the lexically stressed syllables become potential loci of accentuation, usually associated

TEIL 1	PLAY BUTTON	TIMES PLAYED	SYLL. 1	BND	SYLL. 2	BND	SYLL. 3	BND	SYLL. 4	BND	SYLL. 5	BND	SYLL. 6	WORD IDENTIFIED	WORD IDENTIFIED
	<a href="#">PLAY</a>		LISH <input type="checkbox"/>	<input checked="" type="checkbox"/>	PO <input type="checkbox"/>	<input type="checkbox"/>	PU <input type="checkbox"/>	<input type="checkbox"/>	LER <input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	BE <input type="checkbox"/>	<input type="checkbox"/>	LE <input checked="" type="checkbox"/>	POPULÄR	
	<a href="#">PLAY</a>		ER <input type="checkbox"/>	<input type="checkbox"/>	KANT <input type="checkbox"/>	<input checked="" type="checkbox"/>	SHLAK <input checked="" type="checkbox"/>	<input type="checkbox"/>	FER <input type="checkbox"/>	<input type="checkbox"/>	TICH <input type="checkbox"/>	<input checked="" type="checkbox"/>	RE <input type="checkbox"/>	SCHLAGFERTIG	
1.	<a href="#">PLAY</a>		TER <input type="checkbox"/>	<input type="checkbox"/>	KYLT <input type="checkbox"/>	<input type="checkbox"/>	Y <input type="checkbox"/>	<input type="checkbox"/>	BER <input type="checkbox"/>	<input type="checkbox"/>	LEKT <input type="checkbox"/>	<input type="checkbox"/>	UN <input type="checkbox"/>		
2.	<a href="#">PLAY</a>		BE <input type="checkbox"/>	<input type="checkbox"/>	FOL <input type="checkbox"/>	<input type="checkbox"/>	ENT <input type="checkbox"/>	<input type="checkbox"/>	SHI <input type="checkbox"/>	<input type="checkbox"/>	DEN <input type="checkbox"/>	<input type="checkbox"/>	GLAUP <input type="checkbox"/>		
3.	<a href="#">PLAY</a>		BER <input type="checkbox"/>	<input type="checkbox"/>	LEKT <input type="checkbox"/>	<input type="checkbox"/>	FO <input type="checkbox"/>	<input type="checkbox"/>	TO <input type="checkbox"/>	<input type="checkbox"/>	GEN <input type="checkbox"/>	<input type="checkbox"/>	KOM <input type="checkbox"/>		
4.	<a href="#">PLAY</a>		DEN <input type="checkbox"/>	<input type="checkbox"/>	UN <input type="checkbox"/>	<input type="checkbox"/>	TER <input type="checkbox"/>	<input type="checkbox"/>	KYLT <input type="checkbox"/>	<input type="checkbox"/>	FO <input type="checkbox"/>	<input type="checkbox"/>	TO <input type="checkbox"/>		
5.	<a href="#">PLAY</a>		LE <input type="checkbox"/>	<input type="checkbox"/>	GEN <input type="checkbox"/>	<input type="checkbox"/>	GE <input type="checkbox"/>	<input type="checkbox"/>	RI <input type="checkbox"/>	<input type="checkbox"/>	SEN <input type="checkbox"/>	<input type="checkbox"/>	GROS <input type="checkbox"/>		
6.	<a href="#">PLAY</a>		ZEN <input type="checkbox"/>	<input type="checkbox"/>	SHLAK <input type="checkbox"/>	<input type="checkbox"/>	FER <input type="checkbox"/>	<input type="checkbox"/>	TICH <input type="checkbox"/>	<input type="checkbox"/>	LANG <input type="checkbox"/>	<input type="checkbox"/>	VAI <input type="checkbox"/>		
7.	<a href="#">PLAY</a>		FA <input type="checkbox"/>	<input type="checkbox"/>	REN <input type="checkbox"/>	<input type="checkbox"/>	FER <input type="checkbox"/>	<input type="checkbox"/>	LE <input type="checkbox"/>	<input type="checkbox"/>	GEN <input type="checkbox"/>	<input type="checkbox"/>	PRAG <input type="checkbox"/>		

Figure 1: Format of experiment section of questionnaire

with prosodic cues such as F0 transitions and lengthening. Therefore the learner is required to memorize this feature for each word and decode it from the speech stream.

This study concerns the perception of boundaries and accented syllables by native German subjects as compared to Chinese non-speakers and learners of the language at different proficiency levels. We aim to investigate how subjects perform on a task for which they either have to rely – to varying degrees - on their linguistic knowledge as well as the acoustic properties of the stimuli.

Furthermore, we would like to explore the perceptual interrelationship between boundaries and accented syllables, as well as the effect of learning on the judgments as the experiment progresses. This work was inspired by the study of Gilbert, Boucher and Jemel (2011) who showed in a learning experiment that speech chunking in French is performed in rhythm groups. These groups are right-headed, since the final syllable of a content word in French typically bears the prominence. Artificial words straddling a rhythm group boundary are more difficult to recognize than those that terminate at such a boundary. The authors of the current study wish to explore the effects in a language in which accent location is more flexible and decoupled from the word boundary.

## 2 Experiment Design and Stimuli

The stimuli employed in this study are all six-syllable sequences excised from a context of a sentence with the structure “Menschen können A, B, C sein”, (“People can be A, B, C.”) where A, B and C are possible characteristics of people, either adjectives or past participles. These real words possess either two or three syllables, with the lexical accent on the first, second or third syllable, respectively. For each of the conditions we selected 10 unique words none of which contained syllables of the others. From this set of words we constructed

groups of three words A, B and C, where A and C were always three syllables long, and B either contained two or three syllables. Here is an example: “Menschen können be-**le**-sen, **schlag**-fer-tig, **lang**-wei-lig sein.“ (“*People can be erudite, quick-witted, boring*”.) Figure 2 displays a stimulus example marked by a red box with the surrounding carrier phrase.

We intentionally omitted the conjunction ‘and’ before the third item in order not to supply a morphemic marker of boundary. The stimuli were then constructed in such a way, that only the B word was completely preserved in the stimulus whereas only parts of word A and C were present. A stimulus created from the sample sentence could hence be “...lesen, **schlagfertig**, **lang**...”, with the lexical accent syllables set in bold face.

The stimuli were recorded by one male and one female native speaker of German and the target sequences excised using Praat (Boersma & Weenink, 2012). In order for the syllables to be accessible for people without knowledge of German, we used a pseudo-transcription based on German SAMPA in which all non-letter symbols were replaced by other letters or deleted (in the case of the lengthening symbol “:”): LE-ZEN-SHLAK-FER-TICH-LANG. The stimuli, a total of 106 tokens, were then randomized and an MS-Word-based questionnaire was developed which contained a list of the stimuli with PLAY links to the audio files.

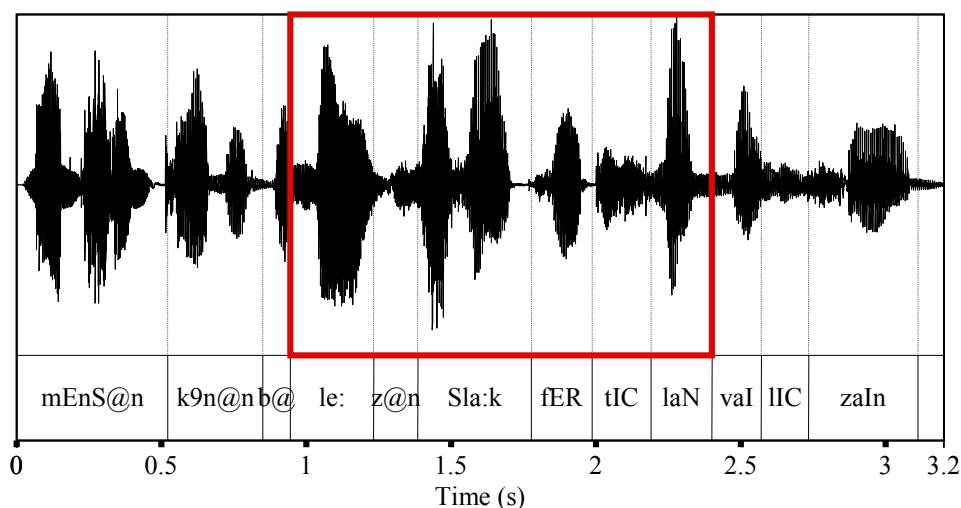


Figure 2: Format of a stimulus. Carrier sentence and six-syllable sequence excised from the context marked by the red box, transcription in German SAMPA.

In the header of the questionnaire we inquired about some personal details such as age and gender, learning history of German, time spent in Germany as well as languages spoken at home. Then the experiment design and its aims were explained, namely, to determine the boundaries between words of German

as well as the accented (“strong”) syllables of these words which had been cut out of context. Participants were advised to play stimuli a maximum number of four times and note down any words they might have perceived. They were also informed that each stimulus contained at least one accented syllable and two boundaries.

The pseudo-transcriptions of the syllables were listed with checkboxes that enabled the selection of each of the syllables/boundaries as shown in the excerpt of the questionnaire displayed in Figure 1 with two illustrative examples in grey and the first seven trials.

Subjects were then asked to work through the examples one by one and make their choices based on their perception. They were also asked to note down how often they played the stimuli and whether they had identified any real words in the sequence. The complete questionnaires were saved in Word format by the participants and later on converted to the readable RTF file format which facilitated access to the states of the checkboxes and hence convenient evaluation of results.

### 3 Results of Analysis

This paper reports results from eight native German listeners, eight Chinese learners after their first year, 15 Chinese learners after their second year, three students after year 5 and eleven Chinese non-speakers of the language. Chinese subjects were all students at the School of Foreign Languages of Tongji University Shanghai. All German participants were naïve listeners without specific linguistic training.

We determined how many of the intended items (boundaries and accented syllables) had been selected by the participants. This measure, however, does not reflect additional selections that were erroneous, that is, insertions. We therefore calculated an error score, in analogy to ASR performance evaluations by defining the error as follows:

$$\text{error} = 100 \times (\text{insertions} + \text{deletions}) / \text{total number of tokens}$$

If we assume that on average two prominences and two boundaries are selected on each trial, chance level would be at 176.7 for the prominences and 175 for the boundaries.

As can be expected (see Table 1), native speakers fare well at identifying word boundaries (95.5% correct) whereas the rate is considerably lower for accented syllables (61.6%). Examination of individual trials shows that there are many cases when even the identification of the word boundaries does not facili-

tate selection of the accented syllable. The 5<sup>th</sup> year students achieve even better results than most of the German natives (96.2%/78.9% correct for boundaries and accents) whereas the ratings are still 69.1% and 46.0% for the non-speakers of German.

Table 1: Ratio correct and insertion/deletion errors

Group	Ratio correct boundaries (mean/s.d.)	Ratio correct prominences (mean/s.d.)	Error boundaries (mean/s.d.)	Error prominences (mean/s.d.)	Number of Subjects (m/f)
Germans	95.5/5.6	61.6/15.0	6.4/8.4	56.7/17.3	2/6
Non-learners	69.1/10.2	46.0/13.5	64.3/19.4	96.2/17.3	3/8
After year 1	80.1/13.1	38.6/22.2	38.1/28.4	81.0/9.9	4/4
After year 2	88.8/7.1	35.1/12.8	20.8/8.8	84.0/11.6	9/6
After year 5	96.2/2.4	78.9/21.3	4.4/2.7	27.0/24.1	0/3

Students on lower levels, that is, after year 1 and 2 improve with respect to word boundaries to 80.1% and 88.8%, respectively. This seems to suggest that their vocabulary and knowledge of word structures expands. However, their performance regarding the accented syllables (38.6% and 35.1%) seems to even deteriorate when compared to the non-learners. This appears like a rather unexpected outcome which can be partly explained by differences in the strategy employed by the subjects. When we divide the mean numbers of insertions by the total number of expected accented syllables, the figure is .38 for non-learners, and only .19 for students after year 2, for instance. In contrast, the number of omissions divided by the total number of accented syllables is .51 for non-learners and .65 after year 2. This means that in general non-learners marked more syllables as accented than the students of German who apparently operated more conservatively and therefore missed a considerable number of accented syllables. We can only speculate about the reasons for this behavior. It might be the case, that the hint in the instructions mentioning *at least* one accented syllable in each stimulus influenced the judgments of the learners.

We examined which factors triggered the selection of an item and compared Germans and Chinese non-learners because these two groups are maximally different with respect to the kind of information they can draw on to access the stimuli. To this effect we calculated the ratio for each item to be selected for the two groups. Table 2 shows the results for the first seven items from Figure 1 for the German subjects, Table 3 those for the Chinese non-learners. As can be seen,

for instance, the syllable ‘LEKT’ was perceived as accented by 86% of the German listeners. The intended correct items are set in bold face.

Table 2: Averaged responses (ratio selected) of German listeners to first seven items.

SYLL.1	BND	SYLL.2	BND	SYLL.3	BND	SYLL.4	BND	SYLL.5	BND	SYLL.6
TER 0.00	0.00	<b>KYLT</b> <b>0.43</b>	<b>0.86</b>	Y 0.00	0.00	BER 0.00	0.00	<b>LEKT</b> <b>0.86</b>	<b>0.86</b>	<b>UN</b> <b>0.00</b>
BE 0.14	0.14	FOL 0.14	0.86	ENT 0.00	0.00	<b>SHI</b> <b>0.86</b>	0.00	DEN 0.00	<b>0.86</b>	<b>GLAUP</b> <b>0.29</b>
BER 0.00	0.00	<b>LEKT</b> <b>0.57</b>	<b>1.00</b>	FO 0.00	0.00	TO 0.00	0.00	<b>GEN</b> <b>0.71</b>	<b>1.00</b>	KOM 0.14
DEN 0.14	<b>1.00</b>	UN 0.00	0.00	TER 0.00	0.00	<b>KYLT</b> <b>1.00</b>	<b>1.00</b>	FO 0.14	0.00	TO 0.00
<b>LE</b> <b>0.57</b>	0.00	GEN 0.00	<b>0.86</b>	GE 0.14	0.00	<b>RI</b> <b>0.71</b>	0.00	SEN 0.00	<b>0.86</b>	<b>GROS</b> <b>0.14</b>
ZEN 0.00	<b>0.86</b>	<b>SHLAK</b> <b>1.00</b>	0.00	FER 0.00	0.00	TICH 0.00	<b>1.00</b>	<b>LANG</b> <b>0.43</b>	0.14	VAI 0.14
<b>FA</b> <b>0.71</b>	0.00	REN 0.14	<b>1.00</b>	FER 0.14	0.00	<b>LE</b> <b>0.57</b>	0.00	GEN 0.29	<b>1.00</b>	PRAG 0.29

Table 3: Averaged responses (ratio selected) of Chinese non-learners to first seven items.

SYLL.1	BND	SYLL.2	BND	SYLL.3	BND	SYLL.4	BND	SYLL.5	BND	SYLL.6
TER 0.00	0.25	<b>KYLT</b> <b>0.50</b>	<b>0.75</b>	Y 0.17	0.08	BER 0.17	0.75	<b>LEKT</b> <b>0.42</b>	<b>0.25</b>	<b>UN</b> <b>0.25</b>
BE 0.00	0.08	FOL 0.50	<b>0.67</b>	ENT 0.33	0.42	<b>SHI</b> <b>0.42</b>	0.25	DEN 0.33	<b>0.58</b>	<b>GLAUP</b> <b>0.25</b>
BER 0.08	0.17	<b>LEKT</b> <b>0.83</b>	<b>0.83</b>	FO 0.08	0.08	TO 0.25	0.67	<b>GEN</b> <b>0.75</b>	<b>0.42</b>	KOM 0.08
DEN 0.50	<b>0.42</b>	UN 0.00	0.33	TER 0.25	0.42	<b>KYLT</b> <b>0.58</b>	<b>0.67</b>	FO 0.17	0.25	TO 0.58
<b>LE</b> <b>0.75</b>	0.33	GEN 0.08	<b>0.67</b>	GE 0.17	0.42	<b>RI</b> <b>0.25</b>	0.33	SEN 0.50	<b>0.58</b>	<b>GROS</b> <b>0.25</b>
ZEN 0.50	<b>0.25</b>	<b>SHLAK</b> <b>0.25</b>	0.75	FER 0.08	0.25	TICH 0.25	<b>0.83</b>	<b>LANG</b> <b>0.75</b>	0.00	VAI 0.00
<b>FA</b> <b>0.33</b>	0.17	REN 0.33	<b>0.83</b>	FER 0.33	0.08	<b>LE</b> <b>0.25</b>	0.17	GEN 0.33	<b>0.83</b>	PRAG 0.42

In the following analysis we do not consider any acoustic measurements, only the structure of the stimuli. Each syllable was classified regarding the following features: (1) lexical stress, (2) vowel length, (3) lexical stress on left/right neighbor, (4) boundary type left/right, (5) number of phones in onset/coda. Boundaries were classified regarding the following features: (1) word boundary, (2) lexical stress on syllable to the left/right. Table 4 shows the results of correlation analysis between the above-mentioned features and the ratio at which a syllable or boundary was selected. Comparison of figures shows that both groups show similar tendencies. The fact that vowel length is correlated with the

ratio at which a syllable was marked as accented is probably as much due to structural properties of German as much as due to vowel length being a prominence lending acoustic feature. In the set of words we selected for this study 68% exhibit lexically stressed syllables with long vowels. An important difference between Germans and non-learners is the effect of adjacent word boundaries on the tendency for a syllable to be perceived as accented. We find negative correlation with the left syllable boundary and positive correlation with the right syllable boundary which is considerably stronger for the non-learners. This means a stronger preference for word-final syllables to be perceived as accented than for word-initial ones. In contrast, the effect of the neighboring syllables' stress status appears stronger in the German subjects, indicating a preference for avoiding stress clash. We also calculated a regression model based on the factors chosen and find that it explains 87% of the variance for the German listeners, but only 68% for the non-listeners.

Table 4: *Factors facilitating the selection of accented syllables and boundaries for Germans and Chinese non-learners.*

Accented Syllables			Boundaries		
Feature	corr. r with ratio selected (Germans)	corr. r with ratio selected (non-learners)	Feature	corr. r with ratio selected (Germans)	corr. r with ratio selected (non-learners)
lexical stress	.76**	.46**	inter-word boundary	.97**	.72**
long vowel	.53**	.38**	stress left	.03(n.s)	.18**
stress left	-.34**	-.25**	stress right	.01(n.s.)	.00 (n.s.)
stress right	-.31**	-.17**			
boundary left	-.13**	-.30**			
boundary right	.11*	.36**			
n onset	.15**	.13**			
n coda	.21**	.24**			
variance explained	.87	.68	variance explained	.97	.74

With respect to the boundaries, the status of the boundary as either being an inter- or intra-word boundary clearly guides the decision whereas the stress status of the adjacent syllables is irrelevant to the judgments of the Germans. In contrast, a stressed syllable to the left of the boundary has a still significant ef-

fect on the judgments of the non-learners. This interrelationship matches the result for the accented syllables.

In terms of the learning effect during the experiment we compared the correctness of results on the first and the second half of stimuli, but did not find any significant differences.

#### **4 Discussion and Conclusions**

Only preliminary results are presented here. We have shown that German listeners, Chinese learners of German on various levels of proficiency and Chinese non-learners of the language perform differently on the accent and boundary identification task. Boundary identification was much more reliable than accent identification for all groups with even the non-learners reaching 69% correct. Chinese learners on an advanced level outperform naïve German listeners.

We analyzed the difference between the judgments of the German subjects and the Chinese non-learners and found - among other results - in the latter a preference to mark syllables adjacent to a word boundary as accented. Work is in progress to look at other possible learning effects, for instance, the effect of correct word identification on subsequent judgment. We also need to consider sub-word units that are legal words of German even if their identification would lead to “non-words” to the left or right. Finally, we will examine the acoustic properties of prominent syllables/boundaries which were easily identified and those which were not.

#### *Acknowledgements*

The second author is supported by the European Structure Fund (ESF). The third author is sponsored by the Innovation Program of Shanghai Municipal Education Commission (12ZS030) for this research work.

#### *Bibliography*

- Boersma, P. & Weenink, D.(2012): Praat: doing phonetics by computer (Version 5.1.26) [Computer program]. Retrieved April 4, 2012, from <http://www.praat.org/>
- Christophe, A., et al. 2003. Discovering words in the continuous speech stream: The role of prosody. *Journal of Phonetics* 31, 585-598.
- Gilbert, A.C., & Boucher, V. J. & Jemel, B. (2011). The Role of Rhythmic Chunking in Speech: Synthesis of Findings and Evidence from Statistical Learning. Proceedings of ICPHS 2011, Hongkong, China.
- Kohler, K. (1977, 1994): Einführung in die Phonetik des Deutschen (Erich Schmidt, Berlin).