Word Recognition in Possible Word Contexts

ADRIANA HANULÍKOVÁ
Humboldt University Berlin
ahanulik@gmail.com

ABSTRACT. The Possible-Word Constraint (PWC; Norris, McQueen, Cutler, and Butterfield 1997) suggests that segmentation of continuous speech operates with a universal constraint that feasible words should contain a vowel. Single consonants, because they do not constitute syllables, are treated as non-viable residues. Two word-spotting experiments are reported that investigate whether the PWC really is a language-universal principle. According to the PWC, Slovak listeners should, just like Germans, be slower at spotting words in single consonant contexts (not feasible words) as compared to syllable contexts (feasible words)—even if single consonants can be words in Slovak. The results confirm the PWC in German but not in Slovak.

1 Introduction

How do listeners find and recognize words in a highly variable and continuous speech signal? The signal does not consist of what we perceive as discrete words, as word boundaries are not marked by pauses.

Previous research has shown that the activation and competition of word candidates are central processes of word recognition (e.g., McClelland and Elman 1986; Norris 1994; for reviews, see Mattys 1997; McQueen 2007). Several acoustic cues to word boundaries have been identified that can effectively modulate the activation of competing candidates and help segment incoming speech (see McQueen and Cutler 2001, for an overview). These cues, such as allophonic details, metrical structure or phonotactic constraints depend on language specific properties. Recent research on word recognition, however, also suggests a universal segmentation principle based on simple information about whether a vowel is present or absent in the speech input.

This universal principle is called the ‘Possible-Word Constraint’ (PWC) (Norris, McQueen, Cutler, and Butterfield 1997; Norris, McQueen, Cutler, Butterfield, and Kearns 2001). According to the PWC, the activation of a word candidate is reduced if between possible word boundaries there is only consonantal material encountered, since not consonants but syllables are viable residues of the input across languages. The first evidence in favour of this claim came from experiments that used the word-spotting task (Cutler and Norris 1988), which was designed to study the segmentation of continuous speech. In this task, listeners hear nonsense words over headphones and respond whenever they detect a real word embedded in these nonsense words. Response times (RTs) and error rates are used as dependent variables. With this task it is possible to compare detection of the same word in different contexts (see McQueen 1996, for further description of this paradigm). In Norris et al.'s (1997) study, listeners were required to spot English words embedded at the beginning or the end of nonsense strings. Norris et al. found that the recognition of a word such as *apple* was significantly faster when the preceding context (contexts underlined in examples) was a syllable (e.g., *vuffapple*) as compared to a single consonant (e.g., *fapple*). None of these contexts are existing English words, but only *vuff* could be a word. Other evidence comes from a study on Sesotho (Cutler, Demuth, and McQueen 2002), where a minimal stand-alone word must contain two syllables. The listeners' performance was significantly faster when a word (e.g., *alafa* 'to prescribe') was embedded in a mono- or a bisyllabic context (e.g., *roalafa* or *pafoalafa*) as compared to a single consonant context (e.g., *halafa*). The crucial result was no significant difference between *pafo* and *ro*, even though only the former is a possible Sesotho word. This study provided strong evidence for the claim that any syllable, but not a consonant is considered as an appropriate parsing unit and that the potential lexical status of words in a specific language is not relevant. The PWC has been replicated in further languages such as Japanese (McQueen, Otake, and Cutler 2001), Dutch (McQueen and Cutler 1998) and Cantonese (Yip 2004), and has been successfully implemented in Shortlist (Norris 1994), a simulation model of word recognition.

However, none of the languages tested so far allows single consonants in its lexical inventory. Berber or Slavic languages, for example, allow single consonants to be words. Such languages would provide a stronger test of the generality of the PWC. The present study thus investigates the universality versus language specificity of the PWC with German and Slovak. Because German does not have words consisting of single consonants, the PWC should apply for German just like for English. Slovak, on the other hand, is a Slavic language where single consonants can be words and thus provides a direct test of the PWC.

Slovak has four prepositions consisting of single consonants such as *k* 'to', *z* 'from', *s* 'with', *v* 'in', each with a voiceless and a voiced positional allophone. These prepositions are proclitic, but they are always separated orthographically by a blank to avoid ambiguity. In order to understand the Slovak phrase *Povedal, že v rane ostal kúsok železa* ('He said that in the wound there was left a piece of iron') the preposition should be segmented from the noun. According to the PWC, however, a word candidate will be disfavoured if there is a single consonant stranded between possible word boundaries. The word candidate *vrane* 'crow' (+Dative inflection) would then win the competition. Thus, the question is whether and how the PWC determines the processing of Slovak consonants during the segmentation of an utterance.

If the PWC is a universal segmentation effect, its predictions should hold for both languages alike. Both Slovak and German listeners should be slower at spotting words in a consonant context as compared to a syllable context. No difference should be obtained between single consonants regardless of their lexical status in Slovak. However, if the PWC is language-specific, Slovak listeners should treat single consonants as viable residues depending on their lexical status. Hence, the recognition of a word in a prepositional context and in a syllable context should both be easier than in a non-prepositional consonant context.

2 Experiment 1

Experiment 1 tested whether German listeners follow the predictions of the PWC and spot a German word (e.g., *Rose* 'rose') in a syllable context (e.g., *suckrose*) faster than in a single consonant context (e.g., *krose*).

Methods

Materials: Forty-eight bisyllabic German words (nouns and verbs) were created. Each word was embedded in three preceding contexts to yield a nonsense string: the consonant /k/ (a preposition only in Slovak, e.g., *krose*), consonants /p, S, t/ (non-prepositions and not possible words in either of the languages, *trose*), and a CVC syllable (possible, but non-existent words in both languages, *suttrose*). Only phonotactically legal consonant clusters with a similar frequency distribution in both languages were chosen. Further, 102 nonsense filler words were created and embedded into the same three preceding contexts. None of the filler words contained real German words. The materials were recorded by a female native speaker of German. The first syllable of the whole string carried the main stress. Three experimental lists were then created with all fillers in each list. Each target appeared only on one list with the type of context counterbalanced over lists. The order of items was pseudo-randomized with the restriction that there was at least one filler item between target-bearing items.

Procedure: Participants were tested individually in a quiet room. They received written instructions that they would hear nonsense strings over headphones. Their task was on each trial to press a button whenever they detected a real word embedded at the end of a nonsense string. They were asked to then say aloud the word they had found. The main experiment started after a short practice session. The presentation of the stimuli and the RT measurements were controlled by NESU (Nijmegen Experiment Set-Up), experimental software developed at the Max Planck Institute for Psycholinguistics in Nijmegen. The stimuli were presented with an interval of 4 s. The RTs were recorded from the stimulus onset, but prior to the analysis were adjusted so as to measure from word offset by subtracting the total sequence duration. Due to a technical problem with the NESU software, which on random trials failed to record participants' responses, no analysis of error rates will be reported in Experiment 1. The available RTs are a subset of the actual response data (approximately 80%).

Participants: Forty native German speakers, students from the Humboldt University in Berlin, volunteered or were paid for participation. They were recruited on the basis of written advertisements. All reported normal hearing.

Results

One subject did not follow the instructions and was excluded from the analysis. One item (*Wesen*) was missed by all but one subject in one condition and was also excluded. The mean RT data and mean error rates are reported in Table 1 and Figure 1. Analyses of Variance (ANOVAs) for both subjects (*F1*) and items (*F2*) were used to test for statistical significance.

	Type of Context		
	C (preposition)	**C (non-preposition)**	**CVC (syllable)**
Example	krose	trose	suckrose
Mean RT	843	859	638
Mean error	24 %	23 %	20 %

Table 1: Experiment 1; Mean RTs (in ms measured from word offset) and mean error rates, C = Consonant, V = Vowel.

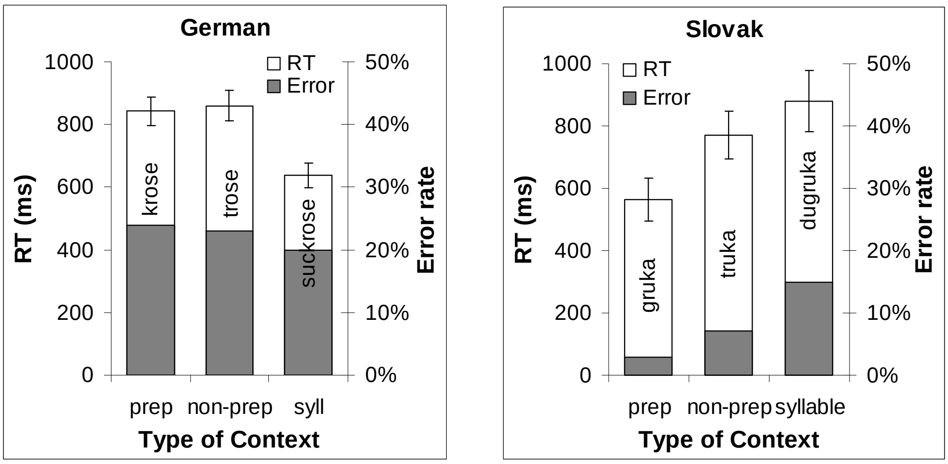

Figure 1: Mean target detection RTs (ms) and mean error rates as a function of type of context, for German and Slovak listeners. Error bars show standard errors. Prep = prepositional consonat, non-prep = non-prepositional consonant, syll = syllable.

The effect of context was significant in the RT analysis for both subjects and items ($F1(2, 76) = 19.87$, $p < .001$; $F2(2, 92) = 15.29$, $p < .001$). Paired comparisons between the contexts revealed that spotting a word in the syllabic context was faster than in both the prepositional context ($t1(38) = 5.09$, $p < .001$; $t2(46) = 5.88$, $p < .001$) and the non-prepositional context ($t1(38) = 6.08$, $p < .001$; $t2(46) = 4.57$, $p < .001$). There was no significant difference in response latencies between the prepositional condition and the non-prepositional condition ($t1(38) = .40$, $p = .69$; $t2(46) = .09$, $p = .93$). To summarize, the results show that spotting a word in the single consonant condition was slower and thus more difficult than in the syllable condition, as predicted by the PWC. This result replicates previous studies and supports the PWC: single consonants are not viable residues of the speech input.

3 Experiment 2

If the PWC is universal and Slovak listeners also follow the PWC, this experiment should replicate the German results. If the PWC is driven by language-specific knowledge, spotting a word in a /g/ context (corresponding to the voiced allophone of the Slovak word *k* 'to') should be equally fast as in a syllable context (possible but non-existent word), and both these conditions should be faster than the non-prepositional context.

Method

Materials: Seventy-two Slovak bisyllabic words (nouns and verbs) were selected. Similar to the German experiment, they were placed in three preceding contexts: prepositions /g, f/ (e.g., *gruka*), non-prepositions /t, S, p/ (e.g., *truka*), and CVC syllables (e.g., *dugruka*). Note that /f/+nouns were not included in the German study, because of the lack of comparable clusters. Although /g, f/ are real words, in combination with verbs and uninflected nouns, the strings as a whole are nonsensical. Verbs were only embedded in /g/ contexts, because /f/ is also a verbal prefix and hence real words would have been formed. Further, 133 fillers were constructed. The materials were recorded by a female native speaker of Slovak. Again, three experimental lists were created.

Procedure: The procedure was identical to that in Experiment 1.

Participants: Nine native speakers of Slovak, students from various universities in Berlin and two employees of the Slovak cultural institute in Berlin, volunteered or received a small payment for participation. All reported normal hearing.

Results

The data of one subject were lost due to a technical failure. Five items were missed by all subjects in one of the conditions and hence excluded (the effects were however not altered by the omission of these items). The mean RTs and the mean error rates for the remaining eight subjects and 67 items are summarised in Table 2 and Figure 1.

There was a main effect of context in both the RT analysis ($F1(2, 14) = 60.13$, $p < .001$; $F2(2, 132) = 16.49$, $p < .001$) and in the error analysis ($F1(2, 14) = 21.99$, $p < .001$; $F2(2, 132) = 7.77$, $p = .001$). Paired *t* tests between the three conditions showed that responses to strings like *gruka* were significantly faster ($t1(7) = 9.97$, $p < .001$, $t2(66) = 3.70$, $p < .001$) than to *truka*, but only marginally more accurate ($t1(7) = 1.82$, $p = .11$, $t2(66) = 1.78$, $p = .08$). Responses to a syllable context like *dugruka* were both slower ($t1(7) = 8.34$, $p < .001$, $t2(66) = 5.97$, $p < .001$) and less accurate ($t1(7) = 6.83$, $p < .001$, $t2(66) = 3.94$, $p < .001$) than to *gruka*. Moreover, *ruka* in *truka* was detected faster ($t1(7) = 4.24$, $p = .004$, $t2(66) = 2.08$, $p = .042$) and more accurately ($t1(7) = 4.63$, $p = .002$, $t2(66) = 2.12$, $p = .038$) than in *dugruka*. None of the participants reported being aware of the presence of prepositions. In summary, the segmentation of a word in the prepositional condition was easier than in the syllable and the non-prepositional condition. This result does not follow the PWC, because word detection should be easier in the syllable context and should be equally hard in both single consonant contexts.

The results instead suggest that segmentation is driven by language-specific knowledge.

	Type of Context		
	C preposition	**C non-preposition**	**CVC syllable**
Example	gruka	truka	dugruka
Mean RT	563	767	880
Mean error	3 %	7 %	15 %

Table 2: Experiment 2; Mean RTs (in ms measured from word offset) and mean error rates, C = Consonant, V = Vowel.

4 General discussion

Previous studies have shown that words enter into an activation and competition process in such a way that consonants are disfavoured as lexical parses regardless of whether they constitute real words in a specific language. This study examined the universality of this Possible-Word Constraint in German and Slovak.

German results replicate previous findings and support the PWC. As predicted, German listeners found it easier to detect a word in a syllabic context than in a single consonant context. The Slovak results however are not consistent with the predictions of the PWC. Target detection was significantly faster in both consonant contexts as compared to a syllable. Moreover, target detection was faster with a preceding preposition than a non-preposition. This suggests that the language-specific knowledge about lexical units modulates the competition process.

The relatively slow responses to a word in the Slovak syllable context also challenge the PWC. Given the theoretical claim, this possible word context should have been the easiest one. The slow responses in this condition are especially striking, as syllables can be words in Slovak. However, an alternative account for this result might be the metrical structure. Slovak, unlike German, has a fixed stress on the first syllable of a word, which could provide a strong cue to locate word boundaries. Thus, the missing canonical main stress on the target (because the stress was on the nonsense syllable) could have slowed down the word recognition as compared to the other conditions where the main stress was on the target. Since little is known about how this regularity in stress assignment is used in Slovak segmentation, further investigation was necessary on this issue and was addressed in a follow-up study. Hanulíková, McQueen, and Mitterer (submitted) have shown that the missing canonical stress indeed contributed to slower reaction times. When the stress information was balanced over conditions, the recognition of a word in the syllable condition was significantly faster as compared to the single non-prepositional consonant context. This result is in line with the PWC. However, in line with the present results, target words were again detected significantly faster in the prepositional condition than in the non-prepositional and the syllable condition. A further lexical decision experiment showed that there were no artefacts in target realisations that would explain the obtained results. Slovak listeners thus appear to use language-specific knowledge about the lexical status of prepositions when segmenting continuous speech.

Due to technical problems with the NESU experimental software in the German experiment, it was necessary to replicate this study too. A follow-up experiment (Hanulíková, 2008) was conducted with a new set of participants. The new study showed the same pattern of results and again replicated the PWC: In German, just like in many other languages, single consonants are not acceptable residues of the input. An additional control lexical decision experiment showed that the word-spotting effects could not be attributed to different acoustic realisations of the targets over conditions.

In summary, the Slovak results support the conclusion that the PWC needs to take language-specific properties into account. Slovak listeners treated single consonants as possible residues of the input, but only when they formed meaningful units. The recognition of target words is most difficult in single consonant contexts that are not possible words in Slovak. Therefore, the PWC needs to be modified and should incorporate language-specific knowledge about minimal meaningful units such as single consonantal words.

Acknowledgments

This work was supported by a grant from the Konrad Adenauer Foundation. For further support I want to thank Rainer Dietrich and James McQueen. I am grateful to all participants of this study, to Katja Kühn and to Marcela Adamíková for recording the material, and to Doug Davidson for proof-reading.

References

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

Cutler, A., K. Demuth, J.M. McQueen (2002). Universality versus language-specificity in listening to running speech. *Psychological Science 13,* 258–262.

Hanulíková, A., J.M. McQueen, H. Mitterer (submitted). Possible words and fixed stress in the segmentation of Slovak speech.

Hanulíková, A. (2008). *Lexical Segmentation in Slovak and German*. PhD Dissertation, Humboldt University Berlin.

Mattys, S.L. (1997). The use of time during lexical processing and segmentation: A review. *Psychonomic Bulletin & Review 4*, 310–329.

McClelland, J.L., J.L. Elman (1986). The TRACE model of speech perception. In *Cognitive Psychology 18*, 1–86.

McQueen, J.M. (1996). Word spotting. *Language and Cognitive Processes 11(6)*, 695–699.

McQueen, J.M. (2007). Eight questions about spoken-word recognition. In G. Gaskell (Ed.), *Oxford Handbook of Psycholinguistics*. Oxford: Oxford University Press.

McQueen, J.M., A. Cutler (1998). Spotting (different types of) words in (different types) of context. *Proceedings of the 5th International Conference on Spoken Language Processing*, 2791–2794. Sydney.

McQueen, J.M., A. Cutler (2001). Spoken word access processes: An introduction. *Language and Cognitive Processes* 16, 469–490.

McQueen, J.M., T. Otake, A. Cutler (2001). Rhythmic cues and possible-word constraints in Japanese speech segmentation. *Journal of Memory and Language 45*, 103–132.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition 52*, 189–234.

Norris, D., J.M. McQueen, A. Cutler, S. Butterfield (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology 34*, 191–243.

Norris, D., J.M. McQueen, A. Cutler, S. Butterfield, R. Kearns (2001). Language-universal constraints on speech segmentation. *Language and Cognitive Processes 16*, 637–660.

Yip, M.C.W. (2004). Possible-word constraint in Cantonese speech segmentation. *Journal of Psycholinguistic Research 33*, 165–173.