Progressive Use of Metrical Cues: A Cross-linguistic Study

Sandrien van Ommen and René Kager
Utrecht Institute of Linguistics, OTS, Utrecht University, Trans 10, Utrecht, The Netherlands

Abstract. Within the framework of a larger project on metrical segmentation this study presents the first results of a cross-linguistic experiment with Dutch (penultimate word stress) and Turkish (word final stress) listeners. Previous studies have shown that listeners interpret stressed or strong (non-reduced) syllables as potential beginnings of words in a.o. English [4], and Dutch [13], [22]. This is interpreted as evidence for the Metrical Segmentation Hypothesis, which predicts that listeners have and use a parsing ability based on edge-aligned stress. However, evidence for a facilitatory effect of right-edge aligned stress is sparse (but see [6]). The current non-word spotting experiment was designed to find out whether listeners can anticipate a word boundary using language-specific stress patterns. The results show that this is partly the case: Dutch listeners are quicker to spot the 'word' when it is preceded by their native penultimate pattern; Turkish listeners are aided by their native final stress pattern as well as by penultimate stress. Turkish listeners, furthermore, make regressive use of metrical cues.

1 Introduction

In running speech, words are not divided by pauses, like spaces in written text, and hence listeners need to develop strategies to segment this stream into meaningful words. Listeners use many different cues to segment speech. Research into this subject has had roughly two different focal points, although efforts have been made to bring the two together (e.g. [12], [21], [10]). The focal points are, on the one hand, sub-lexical cues, which are probabilistically associated with word boundaries. Examples of these cues are phonotactics, phonetic cues and prosodic cues. On the other hand, research has focused on lexically driven segmentation: activation and competition of words.

In the area of sub-lexical cues, much recent interest has gone to metrical segmentation. It has convincingly been shown that listeners are inclined to insert a word boundary before a strong syllable, as predicted by the Metrical Segmentation Strategy (MSS), in English [4], Dutch [13, 21], Finnish [22] and Slovak [5]. This area of research, however, has empirically focused mostly on languages with initial word-stress, leaving many other metrical systems underexposed (but see e.g. [6, 18]). Hence, it remains unclear whether this metrical segmentation is a language-specific or a universal strategy. Furthermore, earlier studies only provided evidence for the use of *regressive* cues, i.e. a stressed syllable leads listeners to infer a word boundary prior to it. The question whether listeners can use native stress to *anticipate* a word boundary has therefore not been answered. The study of Kabak et al. is an exception.

van Ommen S. and Kager R..
Progressive Use of Metrical Cues: A Cross-linguistic Study.
DOI: 10.5220/0004097400740084
In *Proceedings of the 9th International Workshop on Natural Language Processing and Cognitive Science* (NLPCS-2012), pages 74-84
ISBN: 978-989-8565-16-7

It was designed to measure the effects of metrical cues and vowel harmony cues on speech segmentation in French and Turkish. Each trial, an orthographically presented non-word of the form CVCV was presented to the participants, after which an auditorily presented five-syllabic stimulus (all syllables of the form CV) followed. The participants had to react as quickly as possible whenever they spotted the non-word. In the test-items, the target was in final position, preceded by a harmonic or disharmonic prefix with penultimate or final stress. The experiment, thus, tested the progressive use of harmony- and stress cues. The findings were that both French and Turkish listeners used final stress in segmentation, and Turkish listeners additionally benefited from vowel (dis)harmony. This task, like the previously mentioned studies on regressive cues, does not compare languages with different predominant metrical patterns. The current non-word spotting experiment was designed to address this issue. It tests the use of progressive and regressive cues by means of a cross-linguistic comparison of Dutch and Turkish.

Both Dutch and Turkish are languages with canonical right-edge aligned stress. Dutch is a language with predominant penultimate word stress [8], [19], although many exceptions occur. Statistically, due to the high frequency of monosyllabic and bisyllabic words, Dutch is a hybrid of penultimate and initial stress. The Turkish stress system is one with word-final stress. Turkish, too, has classes of exceptions, such as loan words, words containing stress-affecting suffixes, and words following the Sezer stress rule [16], but nevertheless final stress is predominant.

1.1 Predictions

Listeners have knowledge about their native predominant stress pattern. They use this knowledge when they process speech. When this predominant stress pattern is edge-aligned, it can be used to predict word boundaries, i.e. to segment speech. Speech segmentation, then, is facilitated when stress patterns surrounding a word boundary are according to the native predominant pattern. We can distinguish two different kinds of segmentation cues: progressive and regressive cues.

1) Progressive cues: the stress pattern of the preceding context facilitates segmentation; a word boundary is anticipated following a sequence that matches a word-final stress contour.
2) Regressive cues: the stress pattern of the target facilitates segmentation; a word boundary is inferred preceding a sequence that matches a word-initial stress contour.

Because these cues depend on native predominant stress patterns, facilitation should be language-specific: Dutch listeners are aided by their native penultimate pattern and Turkish listeners use their native word-final pattern. On the other hand, the possibility exists that stress is used in a universal way. For example: two subsequent *primary* stresses (clash) have a word boundary in between them. Furthermore, a stressed syllable is more salient than an unstressed syllable, which could be a good reason for it to be a word onset. A long stretch of unstressed syllables (lapse), in the same vein, would be a less salient place for a word boundary. This is especially true for a language with final stress, like Turkish, and a language with statistically frequent initial stress, like Dutch.

2 Method

The word spotting task [11] is widely used in psycholinguistic studies. Listeners detect words embedded in nonce contexts with parsing cues being manipulated. Response latencies and accuracy scores are measured. Because we are dealing with listeners of two different languages, we designed a non-word spotting task based on the task designed by [6], but with important adjustments. The design is attractive, because it can be used cross-linguistically and because it tests progressive metrical cues. However, we wanted our current task to be more like a word-spotting task, and hence we revised Kabak et al.'s task in such a way that participants would first create a lexical entry prior to segmentation. Participants were trained on two auditorily presented non-words, associated with pictures. We presented the non-words auditorily in order to make it possible to test not only the effect of the stress pattern of the prefix (progressive cues), but also that of the target (regressive cues), and their interaction.

2.1 Participants

We tested 38 Dutch and 42 Turkish students at Utrecht University, the Netherlands, and Çukurova University, Turkey, respectively (age 18-26, m= 20.1). They were all monolingual by birth and were tested individually. None of the participants had speech, reading or hearing disorders and all had normal or corrected to normal vision. After the task, each participant filled in a short questionnaire on language contact with other languages than their native language, to confirm their monolingual status. No students had to be excluded on the basis of this questionnaire.

2.2 Material

The materials presented to the participants were five-syllabic nonsense strings, consisting of a three-syllabic prefix with a CVCVCV structure followed by a disyllabic target with a CVCCVC structure. An example of an item, then, would be: /badusudarnam/, of which /darnam/ is the target to be spotted. Each target had one stressed syllable (it is one word-like unit), as did each prefix. Table 1 shows the six experimental conditions according to stress pattern. The items consisted of syllables that are phonotactically legal in both languages. They were controlled for overall syllable frequency and for frequency in stressed and unstressed position. Furthermore,

Table 1. Six different stress-conditions.

condition	prefix			target	
final-final	w	w	s	w	s
final-initial	w	w	s	s	w
penultimate-final	w	s	w	w	s
penultimate-initial	w	s	w	s	w
initial-final	s	w	w	w	s
initial-initial	s	w	w	s	w

w= weak (unstressed), s = strong (stressed) we controlled for possible interfering segmentation cues, such as positional frequency of syllables and vowel harmony. The latter we controlled for by making all strings harmonic according to the harmony-rules in Turkish. Stress-affecting syllables were taken out as well: pre-stressing morphemes [15] and stress-attracting morphemes were not used. Syllable sequences remaining after these control measures were counter-balanced. The items did not contain polysyllabic words in either of the languages. There were two segmentally different targets: one had front vowels and the other back vowels. Each of these items was embedded in 20 segmentally different experimental items. Thus, in total there were 40 segmentally different experimental items. Each individual item was used with each of the different stress patterns, which means that in total there were 6 x 40 = 240 different experimental items, which were divided over 6 lists. The lists did not contain targets which minimally differed in stress, so one list only contained the 'word' with back vowels and initial stress and the 'word' with front vowels and final stress and the other only contained the 'word' with back vowels and final stress and the 'word' with front vowels and initial stress. This was true for experimental items as well as for fillers. Each list consisted of 80 test items as well as 240 filler items, 160 of which did not contain a target and 80 of which contained the target in non-final position, to avoid a bias for position of the target in the string as well as a bias for a positive response.

2.3 Phonetic Resynthesis

The items were recorded integrally by a female native speaker of Spanish. This ensures that there are no phonetic segmentation cues between prefix and target. The choice for a Spanish speaker was made to make the resulting 'language' equally (un)familiar to the Turkish and the Dutch listeners. Each segmentally different item was recorded minimally four times: three times with stress on the first, and three times with stress on the second syllable of the target, the prefix always being flat. To have a natural rhythm on the items, they were recorded in a carrier sentence, in which they received phrase accent on the embedded target. The best token was selected by the researcher. Next, the items were phonetically adjusted through resynthesis to give each condition a different stress pattern while keeping all other phonetic factors constant across conditions. Literature reports that important phonetic correlates for primary stress in Turkish and Dutch are F0, duration and spectral balance (relative amplitude of the higher frequencies) [17], [9]. These factors were adjusted in the rhymes of each syllable, using the program Praat [2]. First, the three vowels of the prefix were made isochronous by giving them the mean duration of the three syllables. Then the stressed syllable was lengthened by a ratio of 1.5. The pitch of the prefix, in turn, was made flat (therefore assuming the pitch of the target) and the stressed syllable received a boost of 8 semitones, with the peak after the first quarter of the overall duration of the vowel. Lastly, the spectral balance of the stressed syllable was adjusted. This latter operation, however, led to a synthetic-sounding result, after which it was decided to give the overall amplitude of the syllable a boost of 8 dB instead. It was important to us to have a natural result after resynthesis. To test the acceptability of the stress patterns we achieved by resynthesis we asked for native speaker judgments of five speakers of each language. They were asked to judge which syllable was stressed and whether the language sounded like 1) their own language, 2) another language 3) a computer language. Nine of the total of ten listeners judged that the language sounded like another language. The remaining listener judged the language to sound like a computer language. Furthermore, the mean accuracy for stress position was far above chance for both language groups.

2.4 Procedure

The experiment was conducted in a quiet room at the university. In Utrecht, the Netherlands, this quiet room was a sound-proof booth at the phonetic lab and in Adana, Turkey this was a regular, but quiet, room. Each participant was randomly assigned to one of six presentation lists. The experiment started with two short animations of creatures called /darnam/ (a red and round creature) and /mernel/ (a yellow triangular creature). During the animations, the participants heard four different tokens of the name of the corresponding creature. This was followed by a static picture of the creature and its name, and a training phase in which the participants had to indicate which creature corresponded to which name. The red creature was always displayed on the left of the screen, the yellow creature on the right. Participants reacted with the button on the corresponding side of the button box. This introduction and training phase is intended to create a lexical entry of the non-word, making the non-word spotting task as close to a word-spotting task as possible. Anecdotally, participants reported that they perceived the language as a proper language and the items as sentences relating to the creatures.

Each item in the practice phase and test phase started with a fixation cross, followed by the auditory presentation of the item. The participant was asked to react as quickly as possible whenever they heard the name of one of the creatures. As soon as the participant recognized one of the words, they had to hit any button of the button box. If and when they did so, a screen appeared with a three-way choice: one creature on the left, one creature on the right and a false alarm button in the middle. Here they chose which name they heard, or used the false alarm button to correct themselves if they had reacted mistakenly. Throughout the test we used an inter stimulus interval of 1 second. When asked afterwards, the participants reported that they thought this was an appropriate speed.

Each participant was tested individually on a laptop in a quiet room. For running and controlling the experiment, the program ZEP was used [20]. The items were presented via Beyerdynamic DT 250 headphones.

3 Results

Statistical analyses were done on the experimental results by use of mixed effects modelling. This technique has as the advantage over ANOVA that it is able to take several random factors into account in one model [1], [14].

The accuracy rate across participants was very high. Overall, in the experimental items and the fillers, the mean accuracy rate is 95.3% (SD = 0.099). Four participants in the Turkish group were excluded on the basis of an accuracy rate of more than 2.5 SD below the mean. The accuracy rate in items containing the target is 95.1% and the accuracy rate in items not containing the target is 95.5%. Because of the high accuracy scores, only response latencies were used as dependent variable. Response latencies were measured from the onset of the target item. In this analysis, only correct answers were included. Response latencies are log-transformed, because of their naturally skewed distribution.

3.1 Model 1: Prefix and Target

To obtain an idea of the random variance in the data, first models with only random effects are run. These models, then, can be compared on the basis of their fitting of the data, and the contribution of the factors to the model. We started with a model explaining variation of the log-transformed latencies by participant only, and subsequently added the segmental structure of the item, segmental structure of the target and handedness (whether the participant is right- or left-hand dominant), respectively. Each subsequent pair of models was compared to see whether the added factor significantly contributed to the model. This model will later be the basis of the final model including fixed factors. Comparing the models showed us that handedness did not significantly contribute to the model. All other factors did, which made the model with the random effects of participant, segmental structure of item and segmental structure of target the best fit. A closer look at this model revealed that, of these random factors, most variance is explained by participant, then segmental structure of the target and segmental structure of the item. To minimize the effect of participant variation in the model, we decided to use the standardized scores of log-transformed response latencies per participant. Outliers of more than 2.5 SD were discarded. A new analysis of models showed that the best fit was a model with segmental structure of the item and segmental structure of the target as random factors. Starting from this model we built a model by subsequently adding the fixed effects of language, prefix and target, and the interaction effects of language with prefix, language with target, prefix with target (see table 1 for the different conditions) and language with prefix with target. Pairwise comparison of each subsequent model showed that the best fitting model was one with the fixed factor of prefix ($F(2)= 10.476$; $p=.000$) and the interaction of language*target ($F(3)= 3.806$; $p=.010$). The other factors did not contribute significantly to the model. A post-hoc analysis with Bonferroni correction of the prefix-effect revealed that overall, participants were significantly slower in the condition with prefix-initial stress ($M=0.090$) than in the condition with penultimate stress ($M= -0.055$; $p= .000$) and the condition with final stress ($M=-0.47$; $p=.000$). In Dutch, the only significant difference in the prefix condition was between prefix initial ($M=0.068$) and penultimate stress ($M=-0.069$, p=..008) and in Turkish the prefix-initial condition was slower than both the penultimate (M=-.0.40, $p=.003$) and the final stress pattern ($M=-.076$, $p=.000$). The interaction effect of language and target proved, in the post-hoc analysis, to be caused by a difference between target-initial ($M=0.053$) and target-final stress ($M=-0.056$; $p=.004$) in Turkish participants, meaning that Turkish participants reacted faster when the target had final stress. In Dutch no such difference existed.

3.2 Model 2: Conditions Collapsed

Because prefix, target and language all had their separate roles in the model, we wanted to see where the different effects originated and whether a combined condition would be a better predictor of current results. We therefore investigated a model with a combination of prefix and target as one variable with six levels (as in table 1). We built this model additively, as described above and compared the fitting of this model with the fitting of model1. The best-fitting model turned out to be the current model: a model with the random factors of segmental structure of the item and segmental structure of the target, and the fixed factor of condition ($F(5)=5.513$; $p=.000$), and the interaction of language*condition ($F(5)=3.186$; $p=.007$). This model was a better fit than model 1. There was no overall effect of language. Post-hoc analyses with Bonferroni correction of condition revealed that, overall, participants were slower in condition SWW[SW] (table 1, row 6) (M=0.150) than in all other conditions but SWW[WS] (not significant): WSW[SW] (M=-0.064; p=.000), WSW[WS] (M=-0.045; p=.002), WWS[SW] (M=-0.047, p=.001) and WWS[WS] (M=-0.047; p=.001). All other comparisons were not significant. Going to the interaction, post-hoc analyses revealed that in Dutch there was a significant effect of condition ($F(5)=2.268$, $p=.046$), but none of the pairwise comparisons turned out significant in the post-hoc. Turkish participants, on the other hand, had slower reaction times in the SWW[SW] condition (M=0.233) than in all other conditions, being SWW[WS] (M=-0.011; p=.003), WSW[SW] (M=-0.019; p=.002), WSW[WS] (M=-0.062; p=.000), WWS[SW] (M=-0.057, p=.000) and WWS[WS] (M=-0.095; p=.000). All other comparisons were not significant. In table 2, the mean standardized logRTs per condition are ordered from low (faster) to high (slower) in both languages.

Table 2. Conditions in order of latency per language. (values are standardized scores).

Dutch		Turkish	
condition	EM means	condition	EM means
WSWSW	−.110	WWSWS	−.095[a]
WWSSW	−.038	WSWWS	−.062[a]
WSWWS	−.028	WWSSW	−.057[a]
WWSWS	.002	WSWSW	−.019[a]
SWWSW	.067	SWWWS	−.011[a]
SWWWS	.070	SWWSW	.233

[a]= is significantly different from the *highest* (slowest) value in Turkish at the .001 level.

Table 2 gives us interesting information. We must recall that in Dutch, the native prefix is WSW, and the native target is SW. In Turkish, the native prefix is WWS and the native target is WS. Universal facilitating effects of metrical patterns would be expected with clash (..S[S..) and native inhibiting effects for both these languages with lapse (..W[W..). We cannot draw firm conclusions from the order as given in table 2, since the significant effects are as given, but it can at least be noted that the most facilitating condition is in fact the native prefix-target sequence of the respective languages (WSW[SW] in Dutch, WWS[WS] in Turkish), and in both languages the facilitating prefix as well as the clash are among the three fastest conditions. The most inhibiting condition is the condition with prefix-initial stress, including lapse (in Dutch) and the least native pattern for Turkish in the Turkish group (SWW[SW]). We will discuss this below.

4 Discussion

The data we analysed first of all showed that the participants had an easy task: the accuracy was very high, with almost no exceptions. The fact that significant effects were found in the data show that not all potential effects have been eliminated by a ceiling effect, but we must be cautious: the differences are very small and a more difficult task might have had, and may in the future have, more distinct results.

The mixed-effects models we built with these data show highly significant effects for all fixed effects, being prefix, target and language, or, as in the second model, condition and language. Model 1 is interesting, because it disentangles progressive and regressive cues. First of all, it brings out an effect of prefix (progressive cue): in both languages, reaction times for both final- and penultimate stress are shorter than for initial stress on the prefix. This is according to the hypothesis, although an interaction effect with language would be expected in the form of facilitation for final stress in Turkish and for penultimate stress in Dutch, as these are the canonical language-specific stress patterns. This interaction-effect was not found. In Dutch, the penultimate pattern was facilitating compared to the initial pattern, but no significant differences with the final pattern were found. No significant differences between the initial and final condition were found either, and it should be noted that the order (by speed) of the prefix conditions is WSW < WWS < SWW. The native penultimate pattern is most facilitating. In Turkish, both the final and the penultimate pattern were facilitating compared to the initial pattern. In this language the order by speed of prefix conditions is WWS < WSW < SWW. Again, the native (final) pattern is most facilitating. However, the significant effects concern the prefix-initial stress pattern. Apparently, it is inhibiting for both languages to move stress away from the right edge of the prefix, increasing the distance between the target and the nearest preceding stressed syllable. These results are not in line with the study of Kabak et al. In this study, there was a significant difference between the prefix condition with penultimate stress and the condition with final stress for Turkish participants. This effect was found in accuracy rates in the disharmonic condition (not in the harmonic condition), and in latencies overall. The study had no condition with prefix-initial stress. It is unclear why this difference in results between this study and the current study exists.

Then, there was an interaction-effect of target and language (regressive cue), showing that, in Turkish, targets with final stress were recognized more quickly than targets with initial stress. A similar facilitation effect of final stress on the target was previously found in a non-word segmentation task with Spanish listeners [3]. No such effect surfaced in the Dutch group, counter to previous findings of a facilitating effect of target-initial stress on word segmentation [13, 21]. Even though these findings do not completely replicate previous studies, we do find cross-linguistic differences. This is seen more clearly in Model 2, collapsing the prefix and target condition. Again, there is an effect of condition, but this time we see that it is to a large extent caused by the SWW[SW] condition. When we, in turn, go to the interaction of language and condition, we see that this effect is largely rooted in the Turkish group. Combining the effect of prefix and language*target, we could say that both the Dutch and the Turkish group appear to be inhibited by the conditions in which the prefix has initial stress, but in the Turkish group, this effect seems to be overridden by a facilitating effect of target, making the SWW[WS] significantly different in latencies from the SWW[SW] condition. Since the hypothesized inhibiting effect of the continuity cue 'lapse' seems to be contradicted through a lack of interaction of prefix- with target stress, there must be some other reason for the encountered delay. This reason may be found in language-specific probabilistic patterns. Up to this point, expectations were based on the canonical stress pattern of the native language. However, both the Dutch and the Turkish language know exceptions to the rule, in the form of lexical stress, pre-stressing morphemes, quantity sensitivity and extrametricality (Dutch; [8]), stress-attracting clitics and Sezer stress (Turkish; [7]). Even though Turkish stress is very regular, it would be interesting to see whether the penultimate stress pattern is statistically common enough to facilitate segmentation compared to the antepenultimate pattern.

Lastly, even though the prefix pattern and the language-target interaction appear to be the main findings of the current study, the mean values of the prefix conditions show an order that is in line with the hypothesis: the native prefix pattern is the most facilitating among the conditions (WWS for Turkish, WSW for Dutch). The mean values of all conditions per language, in turn, confirm that the condition with the native canonical pattern in both the prefix and target (Turkish WWS [WS], Dutch WSW [SW]) is fastest among all other conditions, in both languages. A more challenging task for the participants may create enough room for distinct patterns to appear. Furthermore, research on more languages is needed to confirm whether there is a language-specific interaction between target and prefix, shaking hands to facilitate segmentation.

5 Conclusions

This study shows that non-word spotting is an effective way of testing the use of stress patterns in segmentation. The results show that there is a cross-linguistic effect of stress on segmentation; it is found that Turkish and Dutch listeners make progressive use of metrical cues. Language specific effects have been found, but a more distinct pattern is expected after we tested additional languages.

Acknowledgements

This research was supported by a grant from The Netherlands Organisation for Scientific Research (NWO) within OND1340909 'Parsing and metrical structure: Where phonology meets processing' to René Kager and Wim Zonneveld. We would like to thank Meryem Akcayoglu and Melike Sökmen, who made it possible for us to conduct our experiment at Çukurova University, Adana and who helped us carry out the experiment.

References

1. Baayen, R. H., Davidson, D. J. & Bates, D. M. 2008, "Mixed effects modelling with crossed random effects for subjects and items.", Journal of Memory and Language, vol. 59, no. 4, pp. 390-412.
2. Boersma, P. & Weenink, D. 2011, Praat: Doing phonetics by computer [Computer Program] version 5.3.
3. Cunillera, T., Gomila, A. & Rodriguez-Fornells, A. 2008, "Beneficial effects of word final stress in segmenting a new language: evidence from ERPs", BMC Neuroscience, vol. 9, no. 23.
4. Cutler, A. & Norris, D. 1988, "The Role of Strong Syllables in Segmentation for Lexical Access", Journal of Experimental Psychology: Human Perception and Performance, vol. 14, no. 1, pp. 113-121.
5. Hanulikova, A., McQueen, J. M. & Mitterer, H. 2010, "Possible words and fixed stress in the segmentation of Slovak speech", The Quarterly journal of experimental psychology, vol. 63, no. 3, pp. 555-579.
6. Kabak, B., Maniwa, K. & Kazanina, N. 2010, "Listeners use vowel harmony and word-final stress to spot nonsense words: A study of Turkish and French", Laboratory Phonology, vol. 1, no. 1, pp. 207-224.
7. Kabak, B. & Vogel, I. 2001, "The phonological word and stress assignment in Turkish", Phonology, vol. 18, no. 03, pp. 315.
8. Kager, R. 1989, A metrical theory of stress and destressing in English and Dutch, Foris, Dordrecht.
9. Levi, S. V. 2005, "Acoustic correlates of lexical accent in Turkish", Journal of the International Phonetic Association, vol. 35, no. 01, pp. 73.
10. Mattys, S. L., White, L. & Melhorn, J.F. 2005, "Integration of Multiple Speech Segmentation Cues: A Hierarchical Framework", Journal of Experimental Psychology: General, vol. 134, no. 4, pp. 477-500.
11. McQueen, J. M., Norris, D. & Cutler, A. 1994, "Competition in Spoken Word Recognition: Spotting Words in Other Words", Journal of Experimental Psychology: Learning, Memory, and Cognition, vol. 20, no. 3, pp. 621-638.
12. Norris, D., McQueen, J. M. & Cutler, A. 1995, "Competition and Segmentation in Spoken-Word Recognition", Journal of Experimental Psychology: Learning, Memory, and Cognition, vol. 21, no. 5, pp. 1209-1228.
13. Quené, H. & Koster, M. L. 1998, "Metrical Segmentation in Dutch: Vowel Quality or Stress?", Language & Speech, vol. 41, no. 2, pp. 185-201.
14. Quené, H. & van den Bergh, H. 2008, "Examples of mixed-effects modeling with crossed random effects and with binomial data.", Journal of Memory and Language, vol. 59, no. 4, pp. 413-425.
15. Revithiadou, A., Kaili, H., Prokou, S. & Tiliopoulou, M. 2006, "Turkish accentuation revisited: A compositional approach to Turkish stress", Advances in Turkish Linguistics: Proceedings of the 12th International Conference on Turkish Linguistics, Yağcioğlu, S., A. Cem Değer, Ö. Koşaner & A. Çeltek (eds.), , pp. 37-50.
16. Sezer, E. 1981, "On non-final stress in Turkish", Journal of Turkish Studies, vol. 5, pp. 61-69.
17. Sluijter, A.M.C. & van Heuven, V.J. 1996, "Spectral balance as an acoustic correlate of linguistic stress", The Journal of the Acoustical Society of America, vol. 100, no. 4, pp. 2471-2485.
18. Toro-Soto, J. M., Rodriguez-Fornells, A. & Sebastian-Galles, N. 2007, "Stress placement and word segmentation by Spanish speakers", Psicologica; Vol 28, No 2 (2007), .
19. Trommelen, M. & Zonneveld, W. 1989, Klemtoon en metrische fonologie, Dick Coutinho, Muiderberg.

20. Veenker, T. G. J. 2011, The Zep Experiment Control Application (Version 0.8) [Computer software], Utrecht Institute of Linguistics OTS, Utrecht University.
21. Vroomen, J. & de Gelder, B. 1995, "Metrical Segmentation and Lexical Inhibition in Spoken Word Recognition", Journal of Experimental Psychology: Human Perception and Performance, vol. 21, no. 1, pp. 98-108.
22. Vroomen, J., Tuomainen, J. & de Gelder, B. 1998, "The Roles of Word Stress and Vowel Harmony in Speech Segmentation", Journal of Memory and Language, vol. 38, no. 2, pp. 133-149.