Speaking rate affects the perception of duration as a suprasegmental lexical-stress cue

Eva Reinisch[1], Alexandra Jesse[1], James M. McQueen[1,2]

[1]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

[2]Radboud University Nijmegen, Nijmegen, The Netherlands

Running head: Speaking rate affects suprasegmental lexical stress

accepted: 23 September 2009

in press at Language and Speech

Corresponding author:

Eva Reinisch

Max Planck Institute for Psycholinguistics

P.O. Box 310

6500 AH Nijmegen, The Netherlands

Tel: +31 (0)24 3521 587

Email address: eva.reinisch@mpi.nl

ABSTRACT

Three categorization experiments investigated whether the speaking rate of a preceding sentence influences durational cues to the perception of suprasegmental lexical-stress patterns. Dutch two-syllable word fragments had to be judged as coming from one of two longer words that matched the fragment segmentally but differed in lexical stress placement. Word pairs contrasted primary stress on either the first versus the second syllable or the first versus the third syllable. Duration of the initial or the second syllable of the fragments and rate of the preceding context (fast vs. slow) were manipulated. Listeners used speaking rate to decide about the degree of stress on initial syllables whether the syllables' absolute durations were informative about stress (Experiment 1a) or not (Experiment 1b). Rate effects on the second syllable were visible only when the initial syllable was ambiguous in duration with respect to the preceding rate context (Experiment 2). Absolute second syllable durations contributed little to stress perception (Experiment 3). These results suggest that speaking rate is used to disambiguate words and that rate-modulated stress cues are more important on initial than non-initial syllables. Speaking rate affects perception of suprasegmental information.

(188 words)

Keywords: speaking rate, lexical stress, suprasegmental cues, spoken-word recognition

Information about a spoken word is not presented at once but is rather spread over time. Duration is thus an important perceptual cue in speech comprehension, for example to segmental distinctions. When asked to bring a "bin", the person addressed will bring either a bin or a pin depending on how the duration of the voice onset time (VOT) of the initial phoneme is perceived. Durational cues are, however, not perceived in an absolute fashion but rather relative to the rate at which an utterance is spoken. If the sentence "Can you bring me the bin?" is spoken fast, all segment durations will be shorter. Listeners take this into account and interpret duration relative to speaking rate (e.g., Miller, 1981; 1987). If the word "bin" had the same VOT as spoken at a normal rate but followed a faster spoken carrier sentence, it would sound longer relative to the shortened durational cues in the precursor and be more likely to be interpreted as "pin". But duration does not only distinguish words by cueing the identity of their phonemes. It also provides information about a word's suprasegmental structure, for example, its lexical stress pattern. Among other characteristics, stressed syllables tend to be longer than unstressed syllables. Speaking rate should therefore influence the perception of duration as a cue to stress. The present series of experiments investigated whether this is indeed the case and examined the way speaking rate influences the perception of duration as a suprasegmental cue to stress.

Speaking rate has been shown to affect word recognition by influencing the perception of durational cues to phonemic categories such as the perception of VOT in "bin" vs. "pin" (e.g., Allen & Miller, 2001; Miller & Dexter, 1988; Miller & Liberman, 1979; Summerfield, 1981). Likewise, speaking rate can shift the perception of word boundaries. The perception of a word boundary in pairs like 'great ship' vs. 'grey chip' (Repp, Liberman, Eccardt, & Pesetsky, 1978), for example, is cued by the

duration of the closure before the fricative as well as the duration of the frication noise. Changes in the perceived relation of these durational cues determine whether the plosive-fricative sequence is interpreted as two distinct phonemes spanning the word boundary or as belonging to a single word-initial affricate (Repp et al., 1978). But speaking rate studies so far have focused on durational cues that concern the perception of one or two segments. Here we test whether speaking rate also affects suprasegmental durational cues. The cues we examined specify a word's lexical stress pattern and are distributed over the syllables of a word.

Dutch provides an ideal test case to investigate these issues. Dutch lexical stress is mostly marked suprasegmentally, that is, by changes in duration, pitch, amplitude, and spectral tilt (Cutler, Wales, Cooper, & Janssen, 2007). Although vowel quality changes, where unstressed vowels are reduced to schwa, are common in English, they are rare in Dutch. Unlike spectral cues which can change segmental content (i.e., vowel reduction), duration as a cue to stress is expressed on a segment or syllable but does not change the phonemic content. Although lexical stress can be realized by multiple phonetic cues in Dutch, duration is the most important marker of word-level stress. Whereas pitch cues, for example, depend on sentence intonation (Sluijter & van Heuven, 1996), duration cues are partly preserved even if the sentence focus has been shifted to an unstressed syllable (Sluijter & van Heuven, 1995).

The use of suprasegmental stress information in word recognition can be beneficial in Dutch. There are few Dutch word pairs that are distinguished only by lexical stress (e.g., 'VOORnaam', "first name", and 'voorNAAM, "respectable"; capital letters indicate primary stress). Cutler and van Donselaar (2001) found only 13 semantically unrelated pairs. Nevertheless, taking lexical stress information into account substantially reduces the number of embedded words in the lexicon. A lexical

entry in Dutch contains on average 1.56 embedded words without stress information considered but only 0.74 embedded words with stress information taken into account (Cutler & Pasveer, 2006). Moreover, lexical stress information shifts segmental uniqueness points considerably closer to the beginnings of words (van Heuven & Hagman, 1988). Without stress information taken into account, Dutch words can be uniquely recognized on average after 80% of their phonemes. With stress information considered, this number reduces to 67% of a word's phonemes.

In line with these computational arguments, Dutch listeners indeed use suprasegmental stress information in spoken word recognition (e.g., Cooper, Cutler, & Wales, 2002; Cutler & van Donselaar, 2001; van Donselaar, Koster, & Cutler, 2005; van Leyden & van Heuven, 1996). Dutch listeners distinguish the first two syllables of word pairs like 'Alibi' and 'aLInea' ("alibi" and "paragraph") by means of their lexical stress patterns (e.g., Cutler & van Donselaar, 2001; van Donselaar et al., 2005). Strikingly, Dutch listeners can do this as soon as the suprasegmental information comes available (Reinisch, Jesse, & McQueen, in press) and before the words could be distinguished by segmental information. Since duration is a sufficient cue to recognize the stress patterns of words (Reinisch, Jesse, & McQueen, 2008) speaking rate is expected to influence word recognition by shifting the perception of stress. If the present experiments show that speaking rate influences stress perception, then, given the prior demonstrations of effects of rate on segment identification, the most parsimonious perceptual model would be one with a unitary rate-adjustment mechanism. That is, this mechanism would be concerned not only with adjusting how segmental information is perceived relative to context, but also with adjusting how suprasegmental information is evaluated.

Unlike local durational cues to segmental categories, the lexical stress pattern of multisyllabic words is distributed over syllables. Furthermore, stress is realized to different degrees on the different syllables of a word. A multisyllabic word has one primary stressed syllable and one or more unstressed or secondary stressed syllables. Note that the degrees of stress on the different syllables in a word are not independent of one another. For example, Dutch words with primary stress on the third syllable (e.g., 'kaviAAR', "caviar") have secondary stress on their initial syllable. Different degrees of stress are realized with different durations (Rietveld, Kerkhoff, & Gussenhoven, 2004; Slootweg, 1988). Primary stressed syllables are longer than secondary stressed syllables which in turn are longer than unstressed syllables. Although listeners can distinguish among these syllables (Mattys, 2000), syllables with secondary stress are more difficult to distinguish from primary stressed syllables than unstressed syllables are to distinguish from primary stressed syllables (Reinisch et al., in press). The present study therefore investigates whether speaking rate affects the perception of distinct degrees of stress. Since primary and secondary stress are more confusable than primary stress and no stress, speaking rate may be used to a larger extent for the disambiguation of the more confusable contrast. Moreover, given the distribution of stress patterns over the whole word, we ask here whether speaking rate can affect the perception of stress on more than one syllable in a word.

Although stress is distributed over the syllables of a word, previous research on stress perception has shown that in many cases the initial syllable of a word appears to be sufficient for the perception of stress patterns (van Leyden & van Heuven, 1996). Since the initial syllable of a word is directly adjacent to the rate information of the preceding sentence, it was expected that speaking rate would primarily affect the perception of a word's initial syllable as stressed or unstressed.

This prediction also follows from the view that spoken-word recognition is incrementally optimal: Information is used as soon as it comes available to inform the word-recognition process (Dahan, Magnuson, Tanenhaus, & Hogan, 2001; Norris & McQueen, 2008; Reinisch et al., in press; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Hence, if speaking rate information and the durational information in the first syllable already strongly determine the word's stress pattern, then we expect that rate will primarily affect the perception of that first syllable.

Speaking rate may nevertheless also affect the perception of the second syllable in a word. First, there may not be enough information in the first syllable to fully specify the word’s stress pattern. Information from the second syllable could therefore still make a contribution to optimal word recognition. Second, speaking rate has previously been shown to affect durational cues that are not immediately adjacent to the rate manipulation. The perception of closure duration as a cue to word medial plosives on a 'rabid' - 'rapid' continuum, for example, depends on the rate of the preceding sentence (Port, 1979; Gordon, 1988). Despite the intervening initial syllable, the plosives were perceived more often as voiceless following a fast than a slow carrier sentence. This suggests that the range of the influence of speaking rate could be large enough to affect the perception of stress on the second syllable of a word.

There is, however, a critical difference between stress perception and the perception of segments. As we have already mentioned, the perception of stress cues on the second syllable is not independent of the perception of stress on the initial syllable. The second syllable could be interpreted in relation to the first syllable and speaking rate may thus affect this relation. If both syllables are perceived as long following a fast context, the perceived durational relation between the first two

syllables does not change. Alternatively, however, speaking rate could produce a perceptual chain effect which could cancel or at least diminish the perceived durational difference between the first two syllables, and therefore affect stress perception. That is, following a slow context a stressed initial syllable sounds relatively shorter and, in turn, the second syllable would sound longer relative to the "shortened" initial syllable. This, however, seems unlikely since listeners have no reason to assume a rate change after every syllable. In this study we therefore attempted to control for the dependence of stress on syllables within a word by examining the effect of rate context that is directly adjacent to the second syllable of a word. The initial syllables of the words in the experiment that addresses this issue (Experiment 2) were set to durations that were perceptually ambiguous with regard to stress at a normal speaking rate, and were speeded up or slowed down together with the preceding context. The rate context was thus adjacent to the second syllable while the initial syllable remained ambiguous for stress relative to the preceding context. In this case speaking rate should influence the perception of the second syllable as stressed or unstressed while possible interference from rate effects on the first syllable should be diminished.

The influence of speaking rate on the perception of suprasegmental stress was investigated in a series of two alternative forced choice (2AFC) experiments using the first two syllables of multisyllabic Dutch words. These fragments matched segmentally the beginnings of two longer words which differed in lexical stress placement. 'Alibi' and 'aLInea' have stress on the first vs. the second syllable. There are at least seven other minimal pairs of this type in Dutch (Reinisch et al., in press). 'CAvia' ("guinea pig") and 'kaviAAR' ("caviar") have primary stress on the first vs. the third syllable. There are at least 17 other minimal pairs of this type in Dutch

(Reinisch et al., in press). In addition, 'kaviAAR' has secondary stress on the initial syllable. Stress cues other than duration were neutralized. The listeners' task was to decide as quickly and accurately as possible from which of the two words the fragment (i.e., 'ali' or 'kavi') had been taken. The fragments were presented at the end of a fast or a slow version of a carrier sentence. A fast context should make a syllable sound longer and therefore stressed, a slow context should make it sound shorter and unstressed.

These experiments addressed two aspects of spoken-word recognition. First, they sought to specify the nature of the mechanism(s) which adjust for rate variation. As we noted above, the most parsimonious account of an effect of speaking rate context on the perception of stress patterns is that there is a common mechanism for the evaluation of rate information, one that is concerned with the interpretation of both segmental and suprasegmental information. Second, these experiments tested whether spoken-word recognition is incrementally optimal. If so, speaking rate should have stronger effects on how durational information in a word's initial syllable is interpreted than on its second syllable.

EXPERIMENTS 1a AND 1b

Experiments 1a and 1b investigated whether speaking rate affects the interpretation of stress on the initial syllable of the bisyllabic word fragments. Whereas in Experiment 1a the initial syllable was presented at various durations, in Experiment 1b its duration was kept ambiguous for stress. In this way we asked whether the effect of speaking rate would depend on the informativeness of initial syllable duration.

In Experiment 1a the fragments varied along an initial-syllable duration continuum. The second syllable had an ambiguous duration. It was predicted that

listeners would perceive the initial syllable durations in relation to the preceding rate context. Moreover, despite the ambiguous second syllables, it was predicted that the perceived initial syllable durations should be sufficient for listeners to recognize the stress patterns of the words. Experiment 1b investigated whether speaking rate affects stress perception when the initial syllable was not informative about stress but durational cues were present on the second syllable. Here, the initial syllable was set to an ambiguous duration while the second syllable was varied along a duration continuum. Several outcomes were possible. Speaking rate could affect the perception of initial syllable duration and disambiguate it into a relatively long, stressed syllable following a fast context and a relatively short, unstressed or secondary stressed syllable following a slow context. These perceived initial syllable duration differences could be sufficient for listeners to assign the fragments to the words. If, however, listeners also take the duration of the second syllable into account, then the question is whether listeners would do so in relation to the preceding rate context. On this view listeners should be influenced by initial syllable duration relative to the preceding rate context, and by second syllable duration. Hence, they should give more initial stress responses following a fast context than following a slow context and give more initial stress responses the shorter the second syllable.

Experiment 1b addressed a second question about the perception of different stress patterns. For the fragments from 'Alibi' - 'aLInea', where the second syllable is either short when unstressed or long when stressed, the perception of initial stress should decrease the longer the second syllable. It is unclear, however, how listeners would perceive the stress pattern of the fragment from 'CAvia' - 'kaviAAR' if the second syllable was perceived as long. The second syllables of these word pairs are always unstressed and do not naturally differ in duration. Nevertheless, listeners'

responses to this second syllable continuum can inform us about the mechanisms of stress perception. If the second syllable becomes longer relative to the first syllable, the shorter first syllable might be perceived as not having primary stress. Alternatively, listeners might show a tendency to judge a long syllable on the fragment as having initial stress irrespective of which syllable was long. That is, they could use overall fragment duration rather than initial syllable duration to assign the fragment to the respective word. A fragment with only two short syllables would then be assigned to the word with primary stress outside that fragment, that is, to the word with primary stress on the third syllable.

In summary, Experiments 1a and 1b investigate the effect of speaking rate on the perceived degree of stress on the initial syllable whether initial syllable duration is informative about stress or not. Further, they address the use of second syllable duration in the perception of stress patterns.

Method

Participants. Fourteen Dutch native speakers from the Max Planck Institute participant pool took part for a small payment. They all reported no hearing disorders. A week after completing Experiment 1a they returned for Experiment 1b.

Materials and Design. Two word pairs of at least three syllables length were selected such that both words of a pair overlapped segmentally on their first two syllables but differed in lexical stress placement. 'Alibi' - 'aLInea' ("alibi" - "paragraph") had primary stress on the first vs. the second syllable (1-2 stress contrast). 'CAvia' - 'kaviAAR' ("guinea pig" - "caviar") had primary stress on the first vs. the third syllable (1-3 stress contrast). The word with primary stress on the third syllable ('kaviaar') had secondary stress on the initial syllable. Note that, despite the orthographic difference, the first two syllables of 'cavia' and 'kaviaar' share the same

phonemes. The first two syllables of the words served as fragments that were subjected to the experimental manipulations. For better comparison of durational effects across stress contrasts, word fragments with the same vowels were chosen. Word fragments from multisyllabic words rather than bisyllablic segmental homophones (i.e., words that differ only in stress placement) were used to avoid confounding the effects of duration as a cue to stress with the effects of word-final lengthening (Nooteboom & Doodeman, 1980), and to be able to test different stress contrasts. The words had comparable CELEX lexical frequencies (Baayen, Piepenbrock, & Gulikers, 1995; log-transformed frequencies: Alibi 2.25, aLInea 2.31, CAvia 1.79, kaviAAR 2.21).

The stimuli in all experiments were resynthesized with MBROLA using the diphone inventory of a Dutch female voice. The resynthesis was based on recordings of a Dutch female native speaker who uttered the four words at the end of the carrier sentence 'Klik nog een keer op het woord' ("Click once more on the word"). Word fragments were modeled on the segments of the first two syllables of the words, 'ali' and 'kavi', respectively. They were given a flat pitch contour at the value of the last pitch point in the carrier sentence (190 Hz). The pitch contour in the word fragments was therefore not informative about stress location, neither due to a pitch jump relative to the preceding sentence, nor in its movement within the two syllables. Two versions of the carrier sentence were created to provide a fast and a slow rate context. Segment durations and pitch contours of the sentences were modeled on one token of the recorded sentences. Expansion and compression of speaking rate was implemented linearly by multiplying each of the segment durations in the natural sentence with .66 for a fast version and with 1.33 for a slow version. The same precursor sentences (i.e., fast and slow) were used for both stress contrasts.

In Experiment 1a the duration of the first vowels of the fragments were varied on 13-step duration continua. Since stress mainly affects vowels (Fry, 1955) only durations of vowels were manipulated. Consonants were set to the average of their original durations in the two words of a pair ('ali': [l] 50 ms; 'kavi': [k] 89 ms, [v] 76 ms). Duration values of the vowels were based on a larger data set of similar word pairs described in Reinisch et al. (in press; see Table 1 for a summary of the duration distributions on initial and second vowels in this data set). The lowest values of the continua were set to the values of the appropriate shortest unstressed first vowel in this larger data set. The longest values of the continua approximately matched the appropriate longest stressed first vowel. Only words from the same respective stress contrast as the manipulated fragment were taken into account. The continua spanned durations from 22 ms to 162 ms for 'ali' (1-2 contrast) and from 34 ms to 174 ms for 'kavi' (1-3 contrast). Steps were separated by 11 ms, except for the endpoints which were 15 ms longer or shorter than their adjacent steps. The durations of the second vowels were set to the average duration of second vowels in the same data set (93 ms for 'ali' and 68 ms for 'kavi').

(insert Table 1 about here)

In Experiment 1b, 13-step duration continua were created based on the second vowels of the fragments. Durations were increased in steps of 11 ms from 38 ms to 170 ms for the [i] in 'ali' (1-2 contrast) and from 24 ms to 156 ms for the [i] in 'kavi' (1-3 contrast). Again vowel durations were modeled on the longest and shortest stressed and unstressed vowels in the larger data set. For both fragments the short endpoint of the continuum was set to a duration close to the shortest duration of an unstressed second vowel within word pairs of the same stress contrast. The long endpoints of the continua were established by adding 11 ms 12 times to create the 13-

step continua. Note that the long endpoint for the 1-2 contrast was shorter than that of the longest stressed second vowel in the data set (170 ms vs. 217 ms). This endpoint was chosen because 217 ms was an outlier. For the 1-3 contrast the duration of the short endpoint of the continuum (24 ms) was shorter than the shortest second vowel in the larger data set (33 ms). The long endpoint (156 ms) of the continuum was longer than the longest vowel duration of a second syllable in the data set (112 ms). These values were chosen because words from the 1-3 contrast always have an unstressed, thus short, second syllable. The continuum of the 1-3 contrast should, however, span the same duration range as the continuum for the 1-2 contrast. The vowel durations of the initial syllables were based on the perceptual data from Experiment 1a. They corresponded to the step of each continuum that was perceived as most ambiguous with respect to stress (i.e., the step at which initial and non-initial responses were given about equally often). For 'ali' this corresponded to an [a] of 92 ms and for 'kavi' to an [a] of 104 ms. These values were close to the average durations of the stressed and unstressed first vowels in the larger data set (see Table 1).

Each of the generated 52 fragments (2 Experiments x 2 Stress contrasts x 13 Steps) was combined with both the fast and the slow version of the carrier sentence. Listeners were presented with all stimuli 10 times. Presentation order was randomized separately for each participant within each block such that all combinations of step and rate would be presented once before a repetition occurred. Fast and slow sentences were mixed at random. Presentation was blocked by stress contrast ('ali' and 'kavi' trials) with a switch of contrast after five repetitions. This allowed listeners to answer to one contrast at a time so that they would not respond contrastively. The order of blocks was counterbalanced across participants. In total each participant

responded to 520 stimuli (2 Stress contrasts x 2 Rates x 13 Steps x 10 Repetitions) in each of the experiments.

Procedure. Listeners were tested individually in a sound-attenuated booth. On every trial two response alternatives ('alibi' - 'alinea' or 'cavia' - 'kaviaar' respectively) were presented on a screen with the letters that corresponded to the phonemes of the fragment marked in red. The initially stressed word was always presented on the left side of the screen and corresponded to the left button. After 200 ms listeners heard the carrier sentence immediately followed by the word fragment. Stimuli were presented over headphones at a comfortable listening level. The response alternatives stayed on the screen throughout the trial. The listeners' task was to indicate by button press as quickly and accurately as possible from which of the two words the fragment had been taken. Listeners were not informed that the critical question was about the perception of stress placement. The next trial started 3000 ms after target onset or 500 ms after the participant's response. No feedback was given. Every 65 trials participants were given a short break. The experiment was controlled by NESU experimental software (http://www.mpi.nl/world/tg/experiments/nesu.html).

Analysis. Analyses were run separately for each experiment on responses to the fragments from the 1-2 ('ali') and the 1-3 contrast ('kavi'). Data were analyzed using linear mixed-effect models (Baayen, Davidson, & Bates, 2008) as provided in the lme4 package (Bates & Sarkar, 2007) in R (version 2.8.0; 20-10-2008) with response (i.e., initial stress or not) as dichotomous dependent variable and rate (fast vs. slow), step (numerical values 0 to 12), and block (first vs. second; recoded to the numerical values -1 and 1) as fixed factors. Block was defined for each stress contrast as the first five repetitions of all sentence-fragment combinations vs. the second five repetitions of all sentence-fragment combinations. The variable participant was entered as

random effect in the models. This allowed the intercept to vary by participant with the restriction that the mean of this random variation was zero. A logistic linking function was used to deal with the categorical nature of the dependent variable. The model maps one condition on the intercept (e.g., fast rate at step 0 for the hypothetical block 0) and assigns regression weights for the adjustments required to map from this condition to every other level of each factor. For the categorical factor rate (i.e., fast or slow) the estimated weight gives the mean adjustment of the intercept from fast to slow rate. Step and block were defined as numerical factors. Here, the regression weight has to be multiplied by the numerical value of the factor to adjust the value that has been mapped on the intercept (value 0) to the respective other levels of the factor. A significant effect can be inferred if a regression weight is statistically different from zero. The sign of the weight indicates the direction of the change. In our analyses a positive weight means more initial stress responses, a negative weight means fewer initial stress responses. The analysis started out with a full model that included all three factors and their interactions. Non-significant interactions were eliminated from the model if the simpler model fitted the data better than the model including the interaction. After removing the interactions, non-significant factors were eliminated by the same procedure. During this model-fitting procedure, models were tested against each other by means of log-likelihood ratio tests. Only significant results will be reported.

Results

Experiment 1a. Figure 1 shows categorization data in response to the fragments from the 1-2 and 1-3 contrast in Experiment 1a. For the fragment from the 1-2 stress contrast ('ali') main effects of rate, step, and block were found (b_{rate}=-.43; p<.001; b_{step}=.3, p<.001; b_{block}=-.83, p<.001) as well as an interaction between step and block

($b_{step*block}$=.13, p<.001). Listeners gave more initial stress responses at a fast than at a slow rate context. More initial stress responses were given the longer the first syllable duration and there were more initial stress responses in the first than in the second block. Step had a larger effect in the second block than in the first block, that is, listeners' responses became more categorical in block 2.

(insert Figure 1 about here)

For the fragments from the 1-3 stress contrast ('kavi'), main effects of step and block (b_{step}=.55, p<.001; b_{block}=-.76, p<.001) and an interaction between these two factors ($b_{step*block}$=.11, p<.001) were found. As for 'ali', listeners gave more initial stress responses the longer the first syllable duration and gave more initial stress responses in the first than in the second block. Listeners' responses again became more categorical in the second block. Although there was no main effect of rate, it interacted with step ($b_{step*rate}$=-.08, p<.01). A post hoc test showed that rate affected the longer initial syllable durations (short durations, steps 0-5: b_{step}=.60, p<.001, b_{rate}=-.13, p=.32; long durations, steps 6-12: b_{step}=.35, p<.001, b_{rate}=-.58, p<.001).

Experiment 1b. Figure 2 shows the categorization data for the fragments from the 1-2 and 1-3 stress contrast in Experiment 1b. For the fragment from the 1-2 contrast ('ali') only a main effect of rate was found (b_{rate}= -.82; p<.001). Listeners gave more initial stress responses at the fast than at the slow rate. For the fragment from the 1-3 contrast ('kavi'), rate, step, and block had a significant effect (b_{rate}=-.40; p<.001; b_{step}=.05; p<.001; b_{block}=-.18, p<.01). More initial stress responses were given at a fast than at a slow rate. Step and block also showed a significant interaction ($b_{step*block}$=.03, p<.001). More initial stress responses were given the longer second syllable duration and more initial stress responses were given in the first block. The effect of step was

larger in the second block than in the first block, that is, listeners' responses once again became more categorical later in the experiment.

(insert Figure 2 about here)

Cross-experiment comparison. Additional analyses were carried out in order to compare the effect of first and second syllable durations on the perception of stress. Mean proportions of initial stress responses were calculated for each participant at the endpoints of each duration continuum pooled over rate. Based on the overall proportions at the endpoints, difference scores were calculated for each participant such that the endpoint with the overall lower mean proportion of initial stress responses was subtracted from the endpoint with the overall higher mean proportion of initial stress responses. The difference scores hence indicate the extent of use of syllable duration. Paired t-tests showed that for both stress contrasts the difference in initial stress responses between the endpoints of the duration continua was significantly larger in Experiment 1a than in Experiment 1b (1-2 contrast: $t(13)=4.11$, $p<.001$; 1-3 contrast: $t(13)=8.5$, $p<.001$). Initial syllable duration was used more in the perception of stress location than second syllable duration.

Discussion

Experiments 1a and 1b showed that the speaking rate of the preceding context affects the perception of duration as a suprasegmental cue to lexical stress. The perception of stress location was influenced by the rate whether initial syllable duration was informative about stress (Experiment 1a) or not (Experiment 1b). Following a fast context the initial syllable of the fragment should sound relatively longer, and therefore more stressed than following a slow rate context. Indeed, listeners were more likely to report initial stress when the context was fast than when it was slow. In Experiment 1a listeners considered the steps of the duration continuum

in relation to the speaking rate of the preceding context rather than basing their responses on the absolute duration of the syllables alone. For the 'ali'-fragments rate affected mostly the middle of the duration continuum (steps 3-10; see Figure 1), while for the 'kavi'-fragments the rate effect appears to be stronger for the longer initial vowel durations (steps 5-12). In Experiment 1b the vowel of the initial syllable of each fragment was set to a duration that appeared to be ambiguous for stress in Experiment 1a (i.e., listeners gave approximately 50% initial stress responses at the respective steps of the continua). Listeners in Experiment 1b, however, did not perceive these first syllable durations as completely ambiguous. Rather they interpreted them as long or short in relation to the preceding rate context. Following a fast rate context the initial syllable was perceived as long and therefore stressed; following a slow rate context it was perceived as short and unstressed. The duration of the second syllable (i.e., the different durations of the continuum) contributed little to listeners' responses. Duration variation on the second syllable affected perception only of the 1-3 contrast fragments and this effect was small in comparison to the effect of initial syllable duration in Experiment 1a. The endpoints of the continuum in Experiment 1b received only 45% and 54% initial stress responses as compared to 6% vs. 88% in Experiment 1a. It thus appears that the duration of the initial syllable together with the information from the rate context was largely sufficient to decide about the stress pattern of the word.

The small effect of second syllable duration for the 1-3 contrast, however, provides additional insight in the perception of stress placement. Listeners gave more initial stress responses the longer the second syllable. This suggests that for fragments from the 1-3 contrast a longer duration of either the first or second syllable was perceived as coming from a word with initial syllable stress. That is, listeners were

more likely to perceive initial stress if the fragment contained at least one long syllable. It is only when listeners heard two short syllables that they judged the stress to fall outside the fragment, that is, on the third syllable of the word. Listeners seem to have considered the overall duration of the fragment to decide which word the fragment was taken from. A long fragment was assigned to the initially stressed word whereas a short fragment was assigned to the word with third syllable stress.

Experiments 1a and 1b showed that speaking rate affects the perceived duration of the initial syllable and thereby influences the perception of suprasegmental lexical stress patterns. The second syllable durations of the continuum contributed little to the perception of stress location even if the initial syllable did not carry cues to stress. Nevertheless, speaking rate could in principle affect the perception and use of second syllable duration. Since rate affected the perception of the presumably ambiguous initial syllables in Experiment 1b the use of the duration continuum on the second syllable may not be necessary for rate to affect its perceived duration. Experiment 2 tested whether this was the case. To assess the effect of speaking rate on the second syllable independently from the effect on the initial syllable, the ambiguous initial syllables of Experiment 1b were included in the rate manipulations. That is, the initial syllables were presented at the same rate as the preceding context and therefore were ambiguous with regard to stress within each speaking rate context. Moreover, we asked whether the absence of stress information on the initial syllable would enhance the use of absolute second syllable duration in stress perception. Since the initial syllable duration was not informative about stress and the second syllable was immediately adjacent to the rate context, an effect of second syllable duration as well as an influence of speaking rate on the perception of this duration were expected.

Experiment 2 hence investigated whether speaking rate can affect the perception of suprasegmental lexical stress on the second syllable of a word.

EXPERIMENT 2

Method

Participants. Fourteen participants from the same population as in Experiments 1a and 1b were paid for taking part. None of them had participated in the previous experiments.

Materials, Design, and Procedure. The stimuli from Experiment 1b were manipulated such that the first vowel of each word fragment was speeded up or slowed down at the same rate as the precursor sentence. That is, for the fast context the duration of the first vowel was multiplied by .66; for the slow context it was multiplied by 1.33. The resulting first vowel durations were 61 ms and 122 ms for the 1-2 contrast ('ali') and 69 ms and 138 ms for the 1-3 contrast ('kavi'), respectively. Note that these durations would approximately correspond to steps 3 and 9 of the initial-syllable duration continua (steps from 0 to 12) in Experiment 1a. They were located symmetrically around the middle of the continuum. The duration continua on the second syllable were the same as in Experiment 1b.

The design, experimental setup, task, and analyses were the same as in Experiments 1a and 1b.

Results

Figure 3 shows the proportion of initial stress responses along the continuum for fast and slow rate contexts. For the fragment 'ali' from the 1-2 contrast, main effects of rate and step were found (b_{rate}=.47; $p<.001$; b_{step}=-.03, $p<.01$). Initial stress responses were more frequent after a slow than after a fast context and at shorter second syllable durations. For the fragment from the 1-3 contrast ('kavi') only the

factor rate (b_{rate}= 1.35, p<.001) was significant. More initial stress responses were given following a slow than a fast context.

(insert Figure 3 about here)

Discussion

Listeners again used rate information to perceive the suprasegmental stress pattern of words from both stress contrasts. More initial stress responses were given when the second syllable followed a slow than a fast context. A slow context made the second syllable sound shorter and therefore unstressed. Speaking rate thus affects a syllable that is adjacent to the rate manipulation independent of whether stress cues on that syllable are used or not. The duration of the continuum on the second syllable had little effect on the perception of stress. Unlike in Experiment 1b, however, where a small effect of second syllable duration was found for the 1-3 contrast, here an effect was found for the 1-2 contrast. This suggests that the duration of the second syllable contributes to the perception of primary stress location at least to some extent.

Experiments 1b and 2 support previous findings that the initial syllable is the most important syllable in the perception of Dutch stress patterns (e.g., van Heuven & Hagman, 1988). Note that the absence of any strong effects of the second syllable continuum was not specific to one stress contrast. It is therefore worth asking why listeners did not use much information about the second syllable. Given that in this task listeners gave their response after the offset of the stimuli, it is unlikely that listeners made their decisions about stress location before information from the second syllable came available. But it is possible that listeners might not have perceived the variability in the duration of the fragment-final second vowel. The end of continuous sounds like vowels may be less easily detected when followed by silence than by further speech context. Experiment 3 tested whether this could have

been the case by replicating Experiment 1b with an additional [t] that clearly marked the end of the fragments with its burst.

EXPERIMENT 3

Method

Participants. Fourteen participants who had not taken part in any of the previous experiments received a small payment for their services. None of them reported any hearing problems.

Materials, Design, and Procedure. The sound [t] was added to the stimuli used in Experiment 1b since its burst clearly marks the end of the fragment. The [t] was synthesized with a duration of 100 ms and without intervening silence, so that it was coarticulated with the word-fragments. The sound [t] is not a possible continuation of any of the fragments used here that forms an existing Dutch word. Listeners were informed that they would hear the word fragments 'ali' and 'kavi' to which a [t] had been added (i.e., 'alit' and 'kavit') and that they should decide which word on the screen the fragment was "taken from". The experimenter suggested that the [t] should be ignored. The task was thus similar to gating tasks where fragments are frequently followed by noise (e.g., Smits, Warner, McQueen & Cutler, 2003). Procedure and design were otherwise the same as in the previous experiments.

Results

Figure 4 shows categorization data for the fragments 'ali' and 'kavi' in Experiment 3. For 'ali' (1-2 contrast) a significant main effect of step (b_{step}=-.09, $p<.001$) was found. More initial stress responses were given the shorter the second syllable durations were. Furthermore, the three-way interaction between block, step, and rate ($b_{rate*step*block}$=-.04, $p<.05$) was significant. Subsequent analyses of this interaction for each block revealed a main effect of step in the first block (b_{step}=-.08,

p<.001; more initial stress responses the shorter the second syllable) and an effect of rate and step in the second block (b_{rate}=-.11, p<.001; b_{step}=-.33, p<.001; more initial stress responses the shorter the second syllable and following a fast context). For the fragment from the 1-3 contrast ('kavi') significant main effects of rate, step, and block were found (b_{rate}=-.58, p<.001; b_{step}=.1, p<.001; b_{block}=-.19, p<.01). Listeners gave more initial stress responses at a fast than at a slow rate. Step and block also showed a significant interaction ($b_{step*block}$=.03, p<.001). More initial stress responses were given the longer the second syllable durations were, and more initial stress responses were given in the first block. The effect of step was larger in the second block, that is, responses became more categorical in the course of the experiment.

(insert Figure 4 about here)

Cross-experiment comparison. Experiment 3 was compared to Experiment 1b to test whether the addition of the [t] as a marker of the end of the fragment increased the use of second syllable duration. As for the comparison of Experiments 1a and 1b, difference scores between the proportions of initial stress responses at the endpoints of the duration continua were calculated for each participant pooled over rate. The endpoint with the lower overall mean proportion of initial stress responses was subtracted from the endpoint with the overall higher mean proportion of initial stress responses. Independent sample t-tests showed that in both stress contrasts, second syllable duration was not used more often in the decision about lexical stress location in Experiment 3 (i.e., after the addition of the [t]) than in Experiment 1b (1-2 contrast: t(26)=1.22, p=.23; 1-3 contrast: t(26)=.92, p=.37).

Discussion

Experiment 3 showed that the weak effects of second syllable duration found in Experiment 1b and 2 were not due to the fact that listeners did not perceive the

duration differences in fragment-final position. Although listeners used durational information of the second syllable to decide about the stress pattern of the fragments at least to some extent, they did not do so more when the end of the fragments was marked by the burst of a [t] than when it was followed by silence.

The directions of the effect of second syllable duration for the two stress contrasts were consistent with the effects found in the previous experiments. Whereas for the 1-2 contrast listeners gave more initial stress responses the shorter the second syllable (as in Experiment 2), for the 1-3 contrast more initial stress responses were given at longer second syllable durations (as in Experiment 1b). Although the response pattern for the 1-3 contrast might have been unexpected in Experiment 1b, its replication in Experiment 3 suggests that it did not occur by chance. If the second syllable duration is expected to be uninformative about the stress pattern (i.e., words with primary stress on the first or the third syllable both have a short second syllable) listeners appear to rely on the whole fragment duration. They assigned fragments with a long syllable to the initially stressed word irrespective of whether it was the first or second syllable. Only two short syllables led to the perception of primary stress outside the fragment, that is, stress on the third syllable.

Speaking rate affected the perception of stress such that more initial stress responses were given following a fast than a slow context. For the 1-2 contrast, however, this effect was restricted to the second block. This could have occurred due to task difficulties. Participants reported after the experiments that they found it difficult to ignore the following [t]. With their attention focused on the [t] earlier rate information could have been missed. Note that the [t] could not have provided rate information that attenuated the effect of the preceding context since the duration of the [t] was kept constant across the experiment.

Experiment 3 indicates that listeners can use second syllable duration to decide about the stress patterns of the word fragments. The relatively weaker use of second syllable duration in the perception of stress location as compared to the use of initial syllable duration in Experiments 1a and 1b, however, can not be explained by listeners' failure to perceive the end of the second syllable in fragment final position.

GENERAL DISCUSSION

In a series of categorization experiments we examined whether and how the speaking rate of a preceding sentence context influences the perception of durational cues to the lexical stress patterns of Dutch multisyllabic words. Previous research on speaking rate has focused mainly on rate effects on the perception of segments. Dutch lexical stress, however, is almost exclusively marked suprasegmentally. So Dutch provided a good test of whether speaking rate influences the perception of a purely suprasegmental distinction. Moreover, unlike local durational cues to phonemes or word boundaries, lexical stress patterns of multisyllabic words are distributed over more than one syllable. We therefore asked whether speaking rate can affect stress perception on initial and non-initial syllables and how speaking rate affects the perception of different degrees of stress in different stress patterns (i.e., 1-2 and 1-3 contrasts).

Results showed that speaking rate affected the perception of suprasegmental stress on initial syllables of Dutch word fragments from both tested stress contrasts (the 1-2 contrast and the 1-3 contrast). Since all phonetic information is used as soon as it comes available (Reinisch et al., in press), speaking rate was expected to immediately exhibit its effect on the perception of initial syllable stress. Following a fast context sentence the initial syllable should be perceived as longer and therefore more stressed than following a slow context sentence. Experiments 1a and 1b showed

that speaking rate is not only used to disambiguate syllables that are ambiguous for stress. Speaking rate also contributes to the perception of syllables that could have been assigned to stress patterns on the basis of their absolute durations. Furthermore, similar effects for both stress contrasts showed that speaking rate was used to distinguish different degrees of stress. Primary stressed syllables could be distinguished from unstressed syllables (1-2 contrast) as well as from secondary stressed syllables (1-3 contrast).

Speaking rate also affected the perception of stress on the second syllable. This is remarkable since listeners hardly used the duration continuum on the second syllable to decide about the fragments' stress patterns. As pointed out in the introduction, the stress status of a syllable in a word is not independent of the presence of stress on other syllables. We therefore restricted our investigation to immediately adjacent rate context on the second syllable. That is, the initial syllables of the fragments were made ambiguous for stress and manipulated along with the rate contexts. Experiment 2 showed that despite listeners' lack of attention to the second syllable duration continuum, speaking rate influenced and thus disambiguated second syllable durations. This effect was similar to what was observed with ambiguous initial syllable durations in Experiment 1b. There is, however, an alternative explanation for this rate effect on the second syllable. Rather than using speaking rate to decide whether the second syllable was stressed or unstressed, listeners could have relied on the duration of the initial syllable to decide about the fragments' stress patterns. Note that if the initial syllable was included in a fast context it was short whereas in a slow context it was long. Listeners might not have perceived the initial syllable as included in the rate context and thus not as ambiguous for stress. This use

of initial syllable duration might have led listeners to neglect stress information on the second syllable.

In addition to the main question about speaking rate, the investigation of the 1-3 stress contrast also addressed how listeners perceive stress cues on different syllables. The second syllable of both words from the 1-3 stress contrast is always unstressed, hence expected to be always short. If fragments from the 1-3 contrast were presented with long second syllables, listeners surprisingly but consistently (in Experiments 1b and 3) reported word initial stress. They appeared to use overall fragment durations rather than second syllable durations to interpret the stress patterns. This finding is novel and could not have been found in previous studies on lexical stress perception since these used mostly bisyllabic words (i.e., only stressed vs. unstressed syllables) and, in addition, varied initial and second syllable duration at the same time (e.g., Fry, 1955; van Heuven & Menert, 1996).

Even though the second syllable duration contributed to stress perception to some extent in both stress contrasts, the results emphasized the importance of the initial syllable in stress perception. Experiment 3 thereby replicated Experiment 1b by demonstrating that it was not the failure to perceive the end of the fragment that led to the rare use of second syllable duration. Whenever the duration of the initial syllable of the fragment could be interpreted in its degree of stress, the second syllable duration continuum was given little attention. This is in line with previous findings which also showed that the initial syllable of a word is largely sufficient for listeners to use lexical stress in word recognition (e.g., van Leyden & van Heuven, 1996). This suggests that lexical stress patterns are not perceived solely as duration ratios between initial and second syllables. Rather, stress patterns are perceived incrementally as the signal unfolds (see also Reinisch et al., in press). As with segmental information (e.g.,

Dahan et al, 2001), therefore, suprasegmental information is taken into account as soon as it comes available. The current results thus support the view that spoken-word recognition is incrementally optimal (Norris & McQueen, 2008; Tanenhaus et al., 1995), where all sources of information are used as fully as they can be and as early as possible.

Listeners interpret suprasegmental stress patterns relative to speaking rate information in the preceding context in a similar fashion to how they deal with rate information in making segmental distinctions. Given this similarity between segmental and suprasegmental processing, it is reasonable to assume that there is a common mechanism for the evaluation of both kinds of cues, both locally (i.e., without context information; Reinisch et al., in press) and in relation to preceding context. Moreover, it is plausible to assume that this mechanism has a prelexical locus. Studies on the effects of speaking rate on segmental distinctions suggest that rate effects occur early during processing (Miller & Dexter, 1988; Newman & Sawusch, 2009; Sawusch & Newman, 2000). When listeners rapidly judge which sound they hear, they treat rate information from a following vowel mostly as fast (Miller & Dexter, 1988). That is, they seem to use the available rate information before they processed the complete stimulus. Importantly, this use of rate information appears to be mandatory (Miller & Dexter, 1988) – listeners seem unable to ignore rate information. In addition, speaking rate is taken into account before stream segregation occurs (Newman & Sawusch, 2009; Sawusch & Newman, 2000). If available, listeners use rate information from speakers other than the target speaker. This early and mandatory use of rate information suggests that the rate adjustment mechanism operates at a prelexical processing stage. We suggest further that this is a

unitary mechansim, that is, one through which speaking rate information is used to modulate the interpretation of segmental and suprasegmental cues.

In summary, the present series of experiments demonstrated that speaking rate context affects the perception of durational cues to the lexical stress patterns of Dutch multisyllabic words. Lexical stress patterns on multisyllabic words are distributed over more than one syllable and speaking rate shifts the perception of these patterns by affecting whatever syllable it is adjacent to. The word's initial syllable, however, is most important for listeners to recognize stress patterns. Listeners based their decisions on initial syllable durations even if the syllable's stress status could only be disambiguated by the preceding rate context. The duration of second syllables was used to a lesser extent. The effect of rate was similar for different stress patterns and therefore independent of the location of primary stress in the word. Speaking rate thus provides important context information in spoken word recognition.

AUTHOR NOTE

Correspondence should be addressed to Eva Reinisch, Max Planck Institute for Psycholinguistics, Postbus 310, 6500 AH Nijmegen, The Netherlands. E-mail: eva.reinisch@mpi.nl

This study was conducted as part of the first author's PhD project, funded by the Max Planck Society. It was presented at the 50th meeting of the Psychonomic Society in November 2009, in Boston, USA. The third author is a member of the Behavioural Science Institute and of the Donders Institute for Brain, Cognition and Behaviour, Centre for Cognition, both at the Radboud University Nijmegen. We thank Holger Mitterer for advice about statistical analyses, and Jan McAllister and an anonymous reviewer for helpful feedback.

REFERENCES

ALLEN, J., & MILLER, J. L. (2001). Contextual influences on the internal structure of phonetic categories: A distinction between lexical status and speaking rate. *Perception & Psychophysics, 63*, 798-810.

BAAYEN, H., DAVIDSON, D. J., & BATES, D. M. (2008). Mixed-effect modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.

BAAYEN, H., PIEPENBROCK, R., & GULIKERS, L. (1995). *The CELEX Lexical Database (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

BATES, D. M., & SARKAR, D. (2007). *lme4: Linear mixed-effects models using S4 classes*, version 0.999375-27.

COOPER, N., CUTLER, A., & WALES, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech, 45*, 207-228.

CUTLER, A., & VAN DONSELAAR, W. (2001). Voornaam is not (really) a homophone: Lexical prosody and lexical access in Dutch. *Language and Speech, 44*, 171-195.

CUTLER, A., & PASVEER, D. (2006). Explaining cross-linguistic differences in effects of lexial stress on spoken-word recognition. In R. Hoffmann & H. Mixdorff (Eds.), *Proceedings of Speech Prosody 2006* (pp. 237-240). Dresden: TUD Press.

CUTLER, A., WALES, R., COOPER, N., & JANSSEN, J. (2007). Dutch listeners' use of suprasegmental cues to English stress. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the XVI International Congress of Phonetic Sciences* (pp. 1913-1916). Dudweiler: Pirrot.

DAHAN, D., MAGNUSON, J. S., TANENHAUS, M. K., & HOGAN, E. M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes, 16*, 507-534.

FRY, D.B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America, 27*, 765-768.

GORDON, P. C. (1988). Induction of rate-dependent processing by coarse-grained aspects of speech. *Perception & Psychophysics, 43*, 137-146.

MATTYS, S. L. (2000) The perception of primary and secondary stress in English. *Perception and Psychophysics, 62*, 253-265.

MILLER, J. L., & DEXTER, E. R. (1988). Effects of speaking rate and lexical status on phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 369-378.

MILLER, J. L., & LIBERMAN, A. M. (1979). Some effects of later-occurring information on the perception of stop consonant and semivowel. *Perception & Psychophysics, 25*, 457-465.

MILLER, J. L. (1981). Effects of speaking rate on segmental distinctions. In P. D. Eimas & J. L. Miller (Eds.), *Perspectives on the Study of Speech.* (pp. 39-74). Hillsdale: Erlbaum Associates.

MILLER, J. L. (1987). Rate-dependent processing in speech perception. In A. W. Ellis (Ed.), *Progress in the Psychology of Language* (Vol. 3, pp. 119-157). London: Erlbaum Associates.

NEWMAN, R. S., & SAWUSCH, J. R. (2009). Perceptual normalization for speaking rate III: Effects of the rate of one voice on perception of another. *Journal of Phonetics, 37*, 46-65.

NOOTEBOOM, S. G., & DOODEMAN, G. J. N. (1980). Production and perception of vowel length in spoken sentences. *Journal of the Acoustical Society of America, 67*, 276-287.

NORRIS, D., & MCQUEEN, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review, 115*, 357-395.

PORT, R. (1979). The influence of tempo on stop closure duration as a cue for voicing and place. *Journal of Phonetics, 7*, 45-56.

REINISCH, E., JESSE, A., & MCQUEEN, J. M. (2008). Lexical stress information modulates the time-course of spoken-word recognition. *Proceedings of Acoustics'08 (CD-ROM)*, (pp. 3183-3188). Paris: Société Française d'Acoustique.

REINISCH, E., JESSE, A., & MCQUEEN, J. M. (in press). Early use of phonetic information in spoken word recognition: Lexical stress drives eye-movements immediately. *The Quarterly Journal of Experimental Psychology*.

REPP, B. H., LIBERMAN, A. M., ECCARDT, T., & PESETSKY, D. (1978). Perceptual integration of acoustic cues for stop, fricative, and affricate manner. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 621-637.

RIETVELD, T., KERKHOFF, J., & GUSSENHOVEN, C. (2004). Word prosodic structure and vowel duration in Dutch. *Journal of Phonetics, 32*, 349-371.

SAWUSCH, J. R., & NEWMAN, R. S. (2000). Perceptual normalization for speaking rate II: Effects of signal discontinuities. *Perception & Psychophysics, 62*, 285-300.

SLOOTWEG, A. (1988). Metrical prominence and syllable duration. In P. Coopmans, & A. Hulk (Eds.), *Linguistics in the Netherlands 1988* (pp. 139–148). Dordrecht: Fortis Publications.

SLUIJTER, A. M. C., & VAN HEUVEN, V. J. (1995). Effects of focus distribution, pitch accent and lexical stress on the temporal organization of syllables in Dutch. *Phonetica, 52*, 71-89.

SLUIJTER, A. M. C., & VAN HEUVEN, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America, 100*, 2471-2485.

SMITS, R., WARNER, N., MCQUEEN, J. M., & CUTLER, A. (2003). Unfolding of phonetic information over time: A database of Dutch diphone perception. *Journal of the Acoustical Society of America, 113*, 563-574.

SUMMERFIELD, Q. (1981). Articulatory rate and perceptual constancy in phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 1074-1095.

TANENHAUS, M. K., SPIVEY-KNOWLTON, M. J., EBERHARD, K. M., & SEDIVY, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632-1634.

VAN DONSELAAR, W., KOSTER, M., & CUTLER, A. (2005). Exploring the role of lexical stress in lexical recognition. *The Quarterly Journal of Experimental Psychology, 58A*, 251-273.

VAN HEUVEN, V. J., & HAGMAN, P. (1988). Lexical statistics and spoken word recognition in Dutch. In P. Coopmans & A. Hulk (Eds.), *Linguistics in the Netherlands 1988* (pp. 59-68). Dordrecht: Fortis.

VAN HEUVEN, V. J., & MENERT, L. (1996). Why stress position bias? *Journal of the Acoustical Society of America, 100*, 2439-2451.

VAN LEYDEN, K., & VAN HEUVEN, V. J. (1996). Lexical stress and spoken word recognition: Dutch vs. English. In M. Den Dikken & C. Cremers (Eds.), *Linguistics in the Netherlands 1996* (pp. 159-170). Amsterdam: John Benjamins.

Table 1. Distributions of initial and second syllable vowel durations (ms) for the different stress contrasts and stress locations in the data set described in Reinisch, Jesse, & McQueen (in press). These values were used as reference points for creating duration continua in the present study. Values for the 1-2 contrast are based on seven word pairs; values for the 1-3 contrast are based on 17 pairs. Durations are given in milliseconds.

		Mean duration		Standard deviation		Minimum		Maximum	
stress contrast	stress location	1st vowel	2nd vowel	1st vowel	2nd vowel	1st vowel	2nd vowel	1st vowel	2nd vowel
1-2	1st syllable	107.1	64.4	47	22	53	40	164	99
	2nd syllable	62.3	122.3	28	49	22	86	101	212
1-3	1st syllable	114.6	69.5	38	22	55	34	177	108
	3rd syllable	73.8	65.8	21	22	34	33	132	112

FIGURE CAPTIONS

Figure 1. Proportion of initial stress responses in Experiment 1a for the fragments from the 1-2 ('ali') and 1-3 ('kavi') stress contrast along the duration continuum on the vowel of the respective first syllable. Second syllables were set to the average duration of stressed and unstressed second syllables of tokens from the respective stress contrast.

Figure 2. Proportion of initial stress responses in Experiment 1b for the fragments from the 1-2 ('ali') and 1-3 ('kavi') stress contrast along the duration continuum on the vowel of the respective second syllable. Initial syllables were set to durations that were perceived as ambiguous for stress in Experiment 1a.

Figure 3. Proportion of initial stress responses in Experiment 2 for the fragments from the 1-2 ('ali') and 1-3 ('kavi') stress contrast along the duration continuum on the vowel of the respective second syllable. Initial syllables were set to durations that were perceived as ambiguous for stress in Experiment 1a and included in the rate manipulation.

Figure 4. Proportion of initial stress responses in Experiment 3 for the fragments from the 1-2 ('ali') and 1-3 ('kavi') stress contrast along the duration continuum on the vowel of the respective second syllable followed by a sound [t]. Initial syllables were set to durations that were perceived as ambiguous for stress in Experiment 1a.

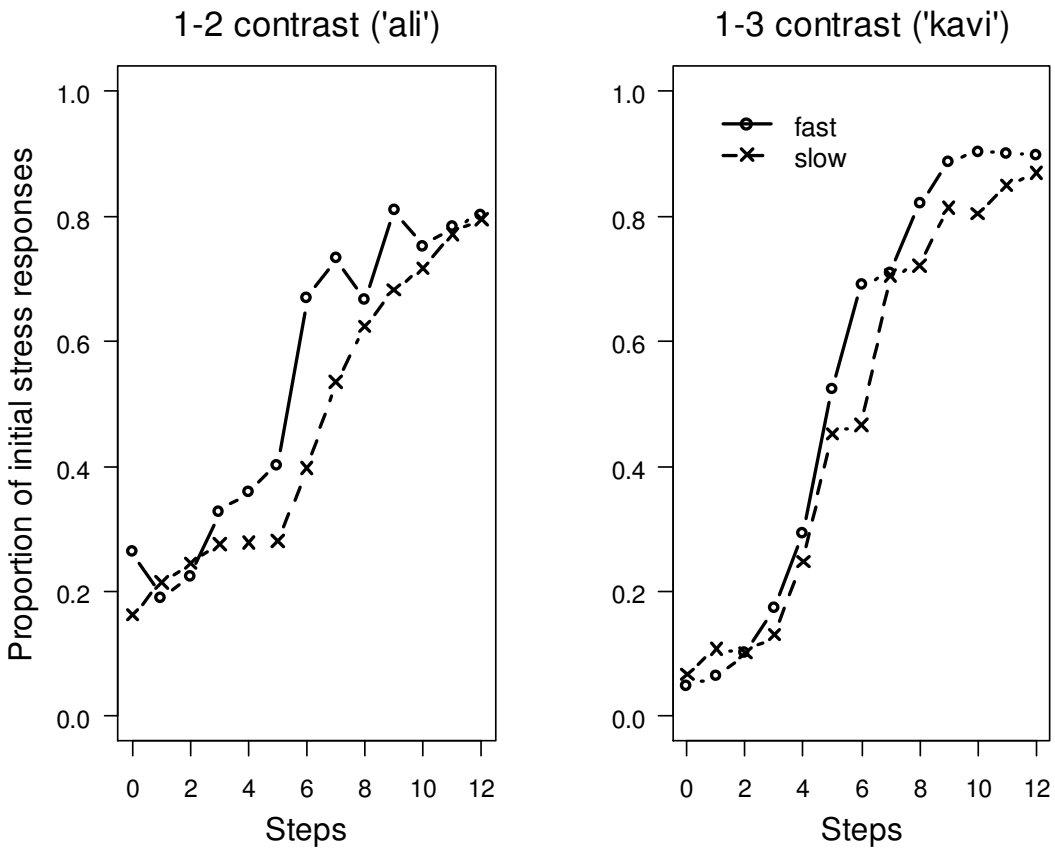
1-2 contrast ('ali')
1-3 contrast ('kavi')
Proportion of initial stress responses
Steps
Steps
fast
slow
0.0
0.2
0.4
0.6
0.8
1.0
0
2
4
6
8
10
12

Figure 1

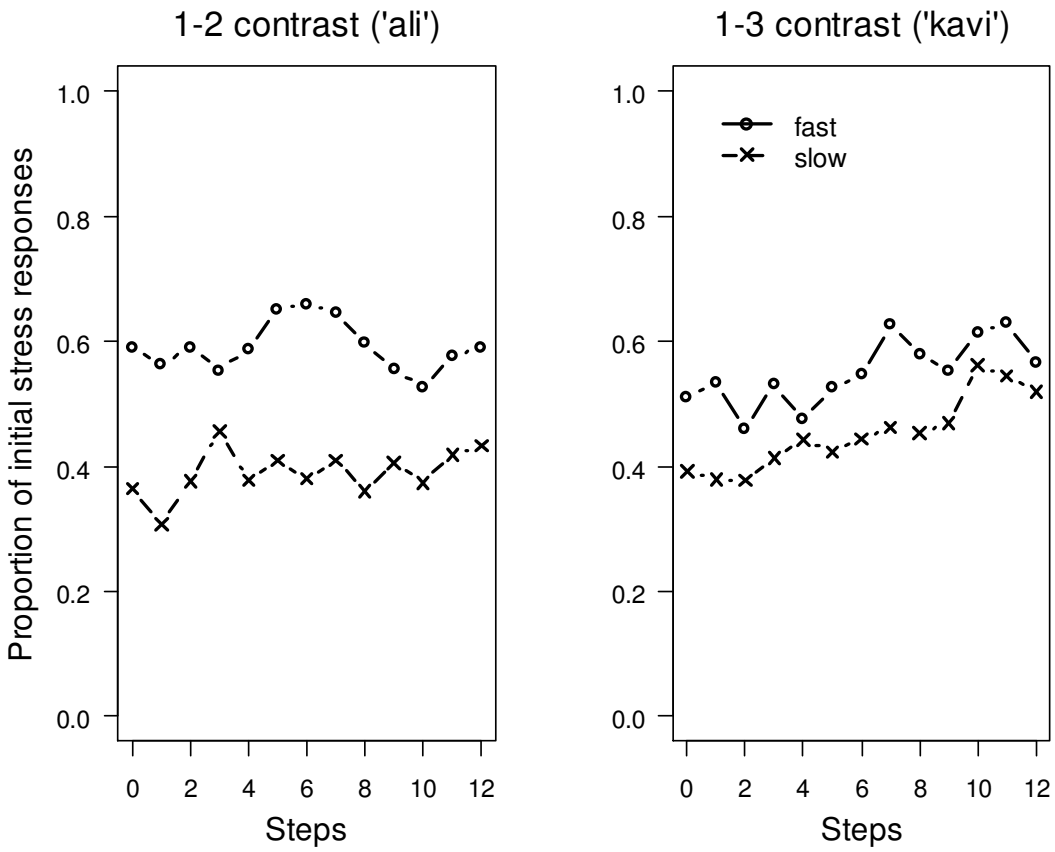
1-2 contrast ('ali')
1-3 contrast ('kavi')
Proportion of initial stress responses
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
12
Steps
fast
slow

Figure 2

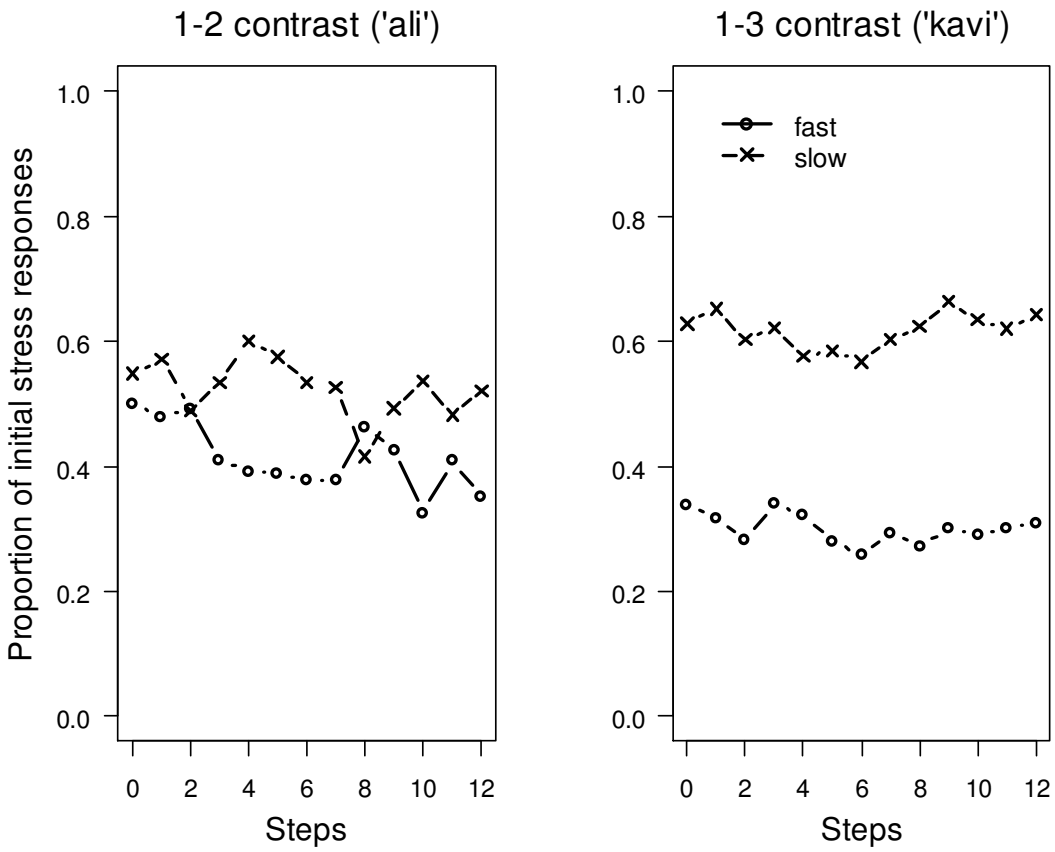
1-2 contrast ('ali')
1-3 contrast ('kavi')
Proportion of initial stress responses
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
12
Steps
fast
slow

Figure 3

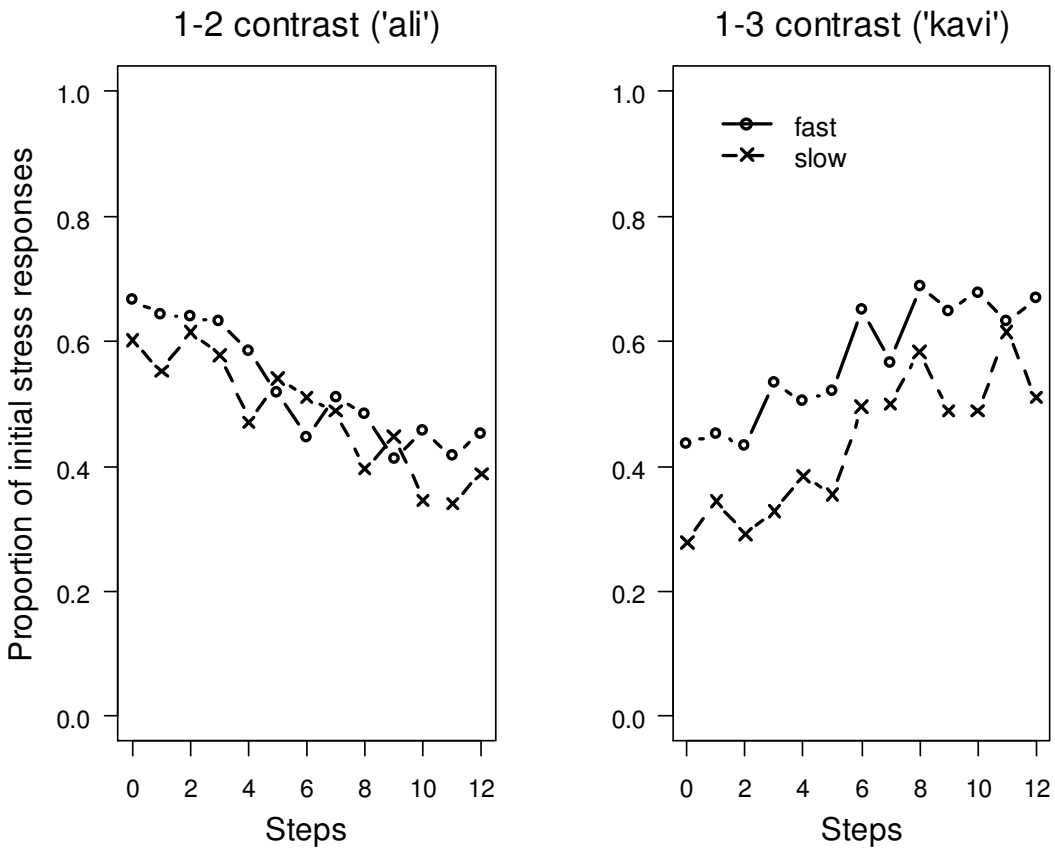
1-2 contrast ('ali')
1-3 contrast ('kavi')
Proportion of initial stress responses
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
12
Steps
fast
slow

Figure 4