The Role of Word-Stress and Intonation in Word Recognition in Dutch 14- and 24-Month-Olds

Article · January 2011

2 authors:

Paula Fikkert
Radboud University
114 PUBLICATIONS  698 CITATIONS

Aoji Chen
Utrecht University
62 PUBLICATIONS  392 CITATIONS

Some of the authors of this publication are also working on these related projects:

www.aojuchen.nl/research View project
The Role of Word-Stress and Intonation in Word Recognition in Dutch 14- and 24-Month-Olds

Paula Fikkert and Aoju Chen

1. Introduction

The ability to segment words from continuous speech is fundamental to language acquisition. In many languages, words have lexical stress. For instance, disyllabic words can have the lexical stress either on the first syllable as in ‘water’, forming the strong-weak (trochaic) pattern, or on the second syllable as in ‘balloon’, forming the weak-strong (iambic) pattern. The strong-weak pattern is the predominant pattern in languages like English, German and Dutch. Adults make use of this information on patterns of lexical stress when segmenting words from continuous speech. For instance, listeners prefer to interpret a strong syllable as the beginning of a word in English (Cutler & Norris 1988) and Dutch (Vroomen & de Gelder 1995; Quené & Koster 1998). Studies on word recognition in infants have shown that English-learning infants as young as 7.5 months are sensitive to the high frequency of the strong-weak pattern and can correctly segment trochaic words from continuous speech, but mistake the stressed syllable of iambic words as the beginning of the word (Jusczyk, Houston & Newsome 1999). Although misstressing words does not appear to impede adults’ word recognition in English (Cutler, Dahan & Van Donselaar 1997), infants’ word recognition seems subject to placement of lexical stress. For instance, Vihman, Nakai, DePaolis & Hallé (2004) found that misstressing familiar trochaic words delayed word recognition in English 11-month-olds, suggesting that either lexical stress is stored in the lexicon of infants in addition to the segmental string or infants expect a word to be trochaic. There is only one study concerning the role of word stress in word recognition in Dutch children. De Bree, Van Alphen, Fikkert & Wijnen (2008) showed that misstressed words were less well recognized than correctly stressed words in Dutch 36-month olds, in particular regarding trochaic words.

However, in everyday speech word production is variable. A word can be produced by speakers of different ages, of different genders, and of different
dialectal backgrounds. Even within a single speaker a word can be produced in different ways, depending on intonation, voice quality and speech rate. These differences inevitably lead to acoustic dissimilarities across tokens of the same word (Singh, White, & Morgan 2008). It is thus essential that young word learners can separate variations that have an impact on the lexical meaning of the word from variations that do not alter the lexical meaning of the word. A main source of within-speaker variation in word production is that speakers realize a word with a different intonation for different communicational purposes. Take for example the word *panda* in the sentence *It’s a panda*. When the sentence is uttered as a statement, *panda* is spoken with a falling pitch accent (i.e. the pitch contour associated with the final word) followed by a low boundary tone (i.e. the pitch level associated with utterance ending). When the sentence is uttered as a question, *panda* is spoken with a rising pitch accent with a high boundary tone.

![Figure 1](image.png)

*Figure 1.* The same sentence produced as a statement and a question

Only very recently, researchers have begun to look at the recognition of words in continuous speech that differ in pitch. Singh et al. (2008) investigated English-learning 7.5-month-olds’ and 9-month-olds’ recognition of familiarized words spoken in different pitch registers (i.e. overall pitch level of an utterance) in passages. The pitch registers of the familiarized words could be either the same as or different from the pitch registers of the words in the familiarization phase. It was found that 7.5-month-olds did not listen to sentences containing the register-mismatched familiarized words longer than sentences containing unfamiliarized words, but they listened to sentences containing the register-matched familiarized words longer than sentences containing unfamiliarized words. This indicated that they recognized the familiarized words only in the absence of a change in pitch register. But 9-month-olds listened longer to sentences containing the familiarized words than the unfamiliarized words regardless of changes in pitch register. This showed that by 9 months, English-learning infants treated variation in pitch register as lexically irrelevant.

Quam and Swingley (2010) examined whether children could recognize a newly learned word (‘deebo’) spoken with a different pitch contour than the tokens heard in the learning phase. In the test phase, children were asked to locate the ‘deebo’ upon hearing the sentence ‘where is the deebo/which one is the deebo? Can you find it?’. The tokens of ‘deebo’ were of three kinds: the same as the tokens in the learning phase, containing a change in pitch contour, or containing a change in the first vowel. The word was spoken either with a weak fall or a (rise-)fall in the testing phase depending on how it was produced.
in the learning phase of each participant. They found that 2.5-year-olds looked above chance at the debo object regardless of change in pitch contour, suggesting that children at this age treated variation in contour shape as lexically irrelevant. Together, these studies show that young children can recognize familiarized and newly learned words independent of changes in pitch register and contour shape by the age of 2.5 years.

To date, no research has been done to examine the interface between word stress and intonation in young children’s word recognition. Singh et al. (2008) used monosyllabic words; Quam and Swingley used tokens of a disyllabic word with the same stress pattern. Considering that word stress and intonation have the same acoustic correlates, variation in intonation would inevitably lead to variation in the acoustics of word stress. Take for example again the word panda in the sentence It’s a panda. In a statement reading of the sentence, panda is spoken with a falling pitch accent. The fall typically starts at the end of the stressed syllable ‘pan’ or at the beginning of the unstressed syllable ‘da’. Perceptually, this gives the high-low impression (in the stressed and unstressed syllable respectively), in accordance with the strong-weak stress pattern. In a question reading of the sentence, panda is spoken with a rising pitch accent with a high boundary tone. The rise typically starts at the end of the stressed syllable, following a low stretch. This generates the low-high impression (in the stressed and unstressed syllable respectively), at odds with the strong-weak stress pattern. How do such changes in the surface form of the words influence infants’ word recognition? Furthermore, how do infants respond to mispronunciations of word stress and inappropriate intonation? Is there an additive effect of word stress and intonation? In the work reviewed above, the change in intonation introduced in the experiment was restricted to a change in the form. However, a change in the form of intonation frequently leads to a change in meaning in everyday speech. Will such a change in meaning influence word recognition? To address exactly these questions, we investigated word recognition in Dutch 14-month-olds and 24-month-olds. As reviewed above, there are a number of studies investigating the role of stress, and a couple of studies investigating the effect of pitch. However, nothing is known about the interface between word stress and intonation in young children’s word recognition.

2. Method

We used a visual fixation task to examine word recognition in children. This task has been regarded as a very sensitive task. Its premise is that better recognition of correctly pronounced words is evidence for well-specified lexical representations (Swingley & Aslin 2000, 2002; Swingley, 2003). If children’s representations of familiar words do not contain all relevant details, minimal mispronunciations will not hinder word recognition and the responses to mispronounced words may not differ from the responses to correct pronunciations. In this procedure, better recognition is assumed to be reflected in a longer looking time to the target picture. In our experiment, familiar words
were used to ensure low processing demands on young children.

2.1. Stimuli

We included four target words in our experiment, two words with the strong-weak stress pattern (hereafter the trochees: varken ‘pig’, schommel ‘swing’), and two words with the weak-strong stress pattern (hereafter the iambic: konijn ‘rabbit’, ballon ‘balloon’). Each target word was paired with a known word starting with the same initial consonant as the target word and having the same determiner. Each target word appeared in two carrier sentences. One was a syntactic question with verb-subject inversion: zie je een .... ‘Do you see a ....’. The other was a syntactic imperative: kijk naar de/het .... ‘Look at the ....’. A 26-year-old female native speaker of Dutch who was known for being good at reading aloud sentences in the required intonation pattern recorded the eight sentences in four conditions in infant-directed speech: (1) correct stress and correct intonation (CSCI), (2) correct stress and incorrect intonation (CSII), (3) incorrect stress and correct intonation (ISCI), and (4) incorrect stress and incorrect intonation (ISII). For the trochees, the incorrect stress pattern was the weak-strong pattern; for the iambic, the incorrect stress pattern was the strong-weak pattern. Regarding intonation, we defined the ‘correct’ intonation as the most typical intonation that was used to utter a sentence, and the ‘incorrect’ intonation as an atypical intonation for the sentence. Thus, the question sentence was produced with a rising pitch accent (L*H) on the target word followed by a high boundary tone (H%) in the CI condition and with a falling accent (H*L) on the target word followed by a low boundary tone (L%) in the II condition. In contrast, the imperative sentence was produced with a falling pitch accent on the target word followed by a low boundary tone in the CI condition and with a rising accent on the target word followed by a high boundary tone in the II condition. The difference in the realization of a pitch accent in these two stress conditions was in how early the fall or the rise started relative to the onset of the syllable that was stressed, as illustrated in Figure 2. Each sentence started with a 50-ms pause. The target word started at about 800 ms after the onset of the sentence across conditions.
The word ‘konijn’ in the ‘zie’ sentence spoken with correct stress and incorrect stress in correct intonation (CSCI, ISCI respectively) and with correct stress and incorrect stress in incorrect intonation (CSII and ISII respectively). The syllable that was stressed in each condition was printed in capitals.

In total there were 32 experimental stimuli: four words (two trochees; two iambics), each occurring in two different sentences (‘zie’ and ‘kijk’ sentences), in four conditions (CSCI, CSII, ISCI, ISII). In addition, four filler sentences were included, consisting of different pairs of known words, which were always produced with correct stress and correct intonation. The fillers contained iambics in the trochee condition, and trochees in the iamb condition. Moreover, the filler sentences were produced with a different verb and sentence meaning than the experimental stimuli.

2.2. Participants

Ninety-two 14-month-olds (43 males and 49 females) and one hundred and one 24-month-olds (56 males and 45 females) participated in the study. An
additional 47 14-month old infants were tested but excluded because of fuzziness or not completing at least half of the trials. An additional 45 24-month olds were excluded for the same reason. The 14-month-olds were beginning word learners, and still developing their inventory of word stress patterns and intonational contours. In contrast, the 24-month-olds had a reasonably large vocabulary (including both trochaic and iambic words), an adult-like inventory of intonation contours (Chen & Fikkert 2007), and knowledge of rise as typical pattern for questions (Leder & Egelston 1982).

2.3. Procedure

The children from each age group were randomly assigned to four participant groups. Each participant group received the stimuli derived from one type of target word in one carrier sentence such that four tests were conducted on each age group: (1) Test A: iambs in ‘zie’ sentences; (2) Test B: trochees in ‘zie’ sentences; (3) Test C: iambs in ‘kijk’ sentences; and (4) Test D: trochees in ‘kijk’ sentences. For each test there were two different orders. Each test consisted of 24 test trials, of which 12 were produced with correct stress and correct intonation (CSCI), 4 trials with ISCI, 4 with CSII and 4 with ISII, and 24 fillers. An attention getter occurred after every six trials to centralize the child’s gaze. In each testing session, the child sat on the parent’s lap facing the centre of the screen of a Tobii 1750 eye-tracker. The parent listened to music via headphones so that he or she could not hear the auditory stimuli. In each trial, the child heard either a ‘zie’ sentence or a ‘kijk’ sentence. All the child had to do was to locate the object on the screen while listening to the sentences. The child’s eye movements were automatically recorded.

3. Analysis

We analyzed the percentage (%) of looking time to the target picture of 86 14-month-olds and 92 24-month-olds in two time frames (TF1 and TF2). TF1 started 360ms after the onset of the target word. This is the point at which the eye movement is assumed to start to reflect the processing of the acoustic information in the target word in young infants (Swingley & Aslin 2000). TF1 lasted for 1 second. TF2 started at the same point as TF 1, but lasted for 2 seconds. The longer TF2 was included as 14-month-olds in general need more time for processing than older children. The % looking time to the target picture in each trial was calculated by dividing the total amount of time that a child spent fixating the target picture by the total amount of time that a child spent fixating the target picture and the distracter picture. On the basis of the literature on word recognition in toddlers, we assumed that a longer looking time meant better recognition of the word.

We conducted four repeated measures ANOVA on the % looking time to the target picture (dependent variable), one for each age group in each time frame. In each ANOVA, there were two within-subject factors, WORD STRESS
4. Results and discussion

With respect to the 14-month-olds, no main effect or interactions were found in TF1. In TF2, the mean % looking time was above chance across conditions, suggesting that the 14-month-olds as a group recognized the words in all conditions. There was a non-significant trend regarding the interaction of Word Stress by Word Type (p = 0.087). In the trochees, the correctness of word stress had little effect: the words were recognized equally well in the CS and IS conditions. This suggests that the 14-month-olds’ recognition of the trochaic words was not impeded by misplacement of word stress, different from the 11-month old English-listening infants in Vihman et al. (2004) and the three-year-olds in De Bree et al. (2007). However, in the iambs IS tended to lead to a longer mean % looking time to the target picture than CS, suggesting that incorrect stress was facilitating word recognition. While it is plausible that children could recognize a misstressed word, it is not very likely that children would recognize a misstressed word better than a correctly stressed word. It might have been the case that 14-month-olds expect words to be trochaic, and hence they listen longer to the expected trochaic pattern. An alternative explanation would be that children look longer to the misstressed iambs because of a ‘surprise’ effect of wrong stress. Considering the fact that the two iambic words in our experiment were fairly frequent in the input and learned early by young children, the 14-month-olds might have been taken by surprise by the misplacement of word stress in ‘konijn’ and ‘ballon’. As a result, they looked longer at the target picture than when the stress was correctly placed. We are, however, not aware of previous studies using this procedure that have found a novelty preference. Our current data on the 14-month-olds are insufficient to settle this issue.

Furthermore, we found a non-significant trend (p = 0.096) regarding the effect of Intonation. That is, correct intonation tended to lead to a longer % looking time than incorrect intonation, suggesting that correct intonation might facilitate word recognition in the 14-month-olds. This result is congruent with the view that correct pronunciation leads to a longer % looking time to the target, and is hence what we expected.

With respect to the 24-month-olds, there was a significant three-way interaction of Word Type by Sentence Type by Word Stress (F = 4.059, p = .047) in TF1. As can be seen in Figure 3 (left panel), in the ‘kijk’ sentences, correct stress triggered a longer mean % looking time than incorrect stress in the trochees, whereas incorrect stress triggered a longer mean % looking time than correct stress in the iambs. Since the incorrectly stressed iambs had the strong-weak stress pattern, just like the correctly stressed trochee, this result suggests
that the strong-weak stress pattern was beneficial in ‘kijk’ sentences (imperatives), regardless of whether it was the correct stress pattern or not. Interestingly, the opposite patterns were observed in the ‘zie’ sentences (Figure 3, right panel). Here, incorrect stress triggered a longer mean % looking time than correct stress in the trochees, whereas correct stress triggered a longer mean % looking time than incorrect stress in the iambs. Since the incorrectly stressed trochees had the weak-strong stress pattern, this result suggests that the weak-strong stress pattern was beneficial in ‘zie’ sentences (questions), regardless of the placement of word stress. Recall that the target word with correct pronunciation was spoken with a falling pitch accent (high-low) in the ‘kijk’ sentences and with a rising pitch accent (low-high) in the ‘zie’ sentences. As mentioned in the introduction, the high-low pitch pattern may trigger the strong-weak impression and the low-high pitch pattern the weak-strong impression. It thus follows that if the perceived stress pattern was in line with the pitch pattern expected on the basis of the sentence-initial verb (‘kijk’ or ‘zie’), this facilitated word recognition in the 24-month-olds, regardless of the underlying stress pattern of the word. This in turn strongly suggests that the 24-month-olds had knowledge of the appropriate pitch accent to be used in questions and imperatives.

**Figure 3.** The % looking time of 24-month-olds in TF1 in imperatives (‘kijk’ sentences) and questions (‘zie’ sentences) for iambs and trochees

In TF2, we found a significant three-way interaction of SENTENCE TYPE by INTONATION by WORD STRESS ($F = 3.949$, $p = .05$). The mean % looking time of the 24-month-olds in different conditions indicated that in the ‘kijk’ sentences, incorrect stress triggered a longer mean % looking time than correct stress when intonation was correct. The opposite pattern was found when intonation was incorrect. In the ‘zie’ sentences this asymmetry did not occur: incorrect stress triggered a longer mean % looking time than correct stress in both the CI and II conditions. We subsequently analyzed the effect of INTONATION and WORD STRESS in the data obtained from the ‘kijk’ sentences and the data obtained from the ‘zie’ sentences separately. We found a significant interaction of INTONATION
by word stress \((F=6.829, p = 0.012)\) in the ‘kijk’ sentences, but neither main effects nor significant interactions in the ‘zie’ sentences.

The results for the ‘kijk’ sentences suggested that correct word stress facilitated word recognition only when intonation was incorrect. When intonation was correct, incorrect stress triggered a longer looking time to the target picture. In line with the possible surprise effect of wrong stress mentioned above, we might speculate that when intonation was correct, the 24-month-olds were struck by the incorrect realization of word stress and looked longer at the target picture as a result of being surprised by the incorrect word stress.

The lack of statistically significant effects of word stress and intonation in the ‘zie’ sentences pointed to a difference between the ‘zie’ sentences and the ‘kijk’ sentences with respect to the choice of pitch accents. Recall that we considered the rising accent as the appropriate pitch accent and the falling accent as the inappropriate pitch accent for the target word in the ‘zie’ sentences. Although the rising accent is the more common pitch accent in questions, questions can be pronounced with a falling pitch accent, as well. In the case of the ‘zie’ sentences, the degree of interrogativity is weakened by the falling accent and the sentence can function like a strong imperative. It was thus possible that in the II condition in the ‘zie’ sentences, some 24-month-olds had the question interpretation of the ‘zie’ sentences and some had the imperative interpretation. Such within-subject variability may account for the lack of main effect of word stress in the II condition. The lack of main effect of word stress in the CI condition might reflect the relatively weak status of a rising accent as the correct intonation in the ‘zie’ sentences, compared to the status of a falling accent as the correct intonation in the ‘kijk’ sentences.

![Figure 4](image.png)

**Figure 4.** The % looking times of 24-month-olds to target words with correct vs. incorrect stress in the correct intonation vs. incorrect intonation in ‘kijk’ sentences
5. Conclusions

Generally, we have found that the mean % looking time was above chance across conditions in both age groups. This indicates that both the 14-month-olds and the 24-month-olds recognized the target words in not only correct stress and correct intonation conditions but also in incorrect stress and/or incorrect intonation conditions. Importantly, the two groups of children differed both in how early in the process of word recognition they showed sensitivity to word stress and intonation and in how exactly the two variables affected their word recognition.

The 14-month-olds appeared to have processed word stress and intonation only in TF2 but showed no sign of integrating these two factors in word recognition. They tended to recognize the target words better when intonation was pragmatically appropriate than when the intonation was pragmatically inappropriate. Further, they seemed to recognize the target words equally well regardless of the appropriateness of word stress in the case of the trochaic words, suggesting robust recognition of trochaic words, but might favour the incorrect stress pattern over the correct stress pattern in the case of the iambic words, which could be due to an inherent bias towards the trochaic pattern or to being surprised by the wrong word stress in the frequent iambic words.

In contrast, the 24-month-olds showed sensitivity to word stress and intonation in TF1. They anticipated the pragmatically appropriate intonation in the target words on the basis of the (semantics of the) sentence-initial verbs, and recognized the target words better when the pronounced stress pattern acoustically accorded with the anticipated intonation pattern: an expected fall facilitated the recognition of correctly stressed trochees and incorrectly stressed iambs; an expected rise facilitated the recognition of correctly stressed iambs and incorrectly stressed trochees. In TF2, they processed the produced stress pattern and the produced intonation in an interactive fashion: correct stress facilitated word recognition only when intonation was incorrect. There was thus no additive effect of correct intonation and correct stress. When the intonation was correct, incorrect stress seemed to trigger a surprise effect, resulting in a longer looking time to the target picture.

Our results thus showed that the pragmatic appropriateness of an intonation pattern had an effect on word recognition in young children, in particular in the 24-month-olds. Crucially, our study provided the first evidence for a two-stage processing: the 24-month-olds processed the anticipated intonation patterns first, and subsequently integrated the produced intonation pattern together with the produced stress pattern in the target word.

References


