FAST SPEECH TIMING IN DUTCH: DURATIONAL CORRELATES OF LEXICAL STRESS AND PITCH ACCENT

Esther Janse, Anke Sennema & Anneke Slis

Utrecht institute of Linguistics OTS, Utrecht University, The Netherlands

ABSTRACT

In this study we investigated the durational correlates of lexical stress and pitch accent at normal and fast speech rate in Dutch. Previous literature on English shows that durations of lexically unstressed vowels are reduced more than stressed vowels when speakers increase their speech rate. We found that the same holds for Dutch, irrespective of whether the unstressed vowel is schwa or a 'ull' vowel. In the same line, we expected that vowels in words without a pitch accent would be shortened relatively more than vowels in words with a pitch accent. This was not the case: if anything, the accented vowels were shortened relatively more than the unaccented vowels. We conclude that duration is an important cue for lexical stress, but not for pitch accent.

1. INTRODUCTION

When speech is time-compressed artificially, compression is normally performed in a linear way. When talking fast, however, speakers do not compress speech segments in a linear way. Duration studies on English have shown, for instance, that when people speak faster, consonant durations are reduced less than vowel durations (Lehiste 1972, Gay 1978, Max & Caruso 1997). Another interesting finding is that durations of lexically unstressed syllables are reduced more, relatively speaking, than durations of stressed syllables (Peterson & Lehiste 1960, Port 1981, Van Santen 1994). As a result, the relative difference in duration between stressed and unstressed syllables increases, thereby making the prosodic pattern more prominent.

This study reports the results of a production experiment, set up to find out whether the effect of increasing speech rate on word-level timing also holds for Dutch. As we are working on the intelligibility of time-compressed speech, the ultimate goal of carrying out the duration study was to test the hypothesis that the intelligibility of time-compressed speech might be improved if its temporal organisation is closer to natural fast speech (cf. Janse 2000). For that purpose, we investigated the differences between the timing of normal and fast speech in Dutch with respect to lexical stress and pitch accent.

For Dutch, we expect that the unstressed syllables are reduced more in duration than stressed syllables as speech rate increases, as was found for English. In particular, unstressed vowels will be shortened more than stressed vowels. Yet, there is a linguistic difference between Dutch and English. In English, unstressed vowels are mostly reduced to schwa at normal speech rate. In Dutch the unstressed syllable may well contain a 'ull' vowel. In order to investigate whether the data found for English might be caused mainly by most unstressed vowels being schwa, we will test the shortening of unstressed schwa and unstressed full vowels, compared to the shortening of stressed vowels. Nevertheless, we expected that increasing speech rate would have a greater effect, relatively, on the unstressed vowel than on the stressed vowel. We did expect, however, that the duration of schwa would be reduced more than that of 'ull' unstressed vowels.

The studies on fast speech timing and lexical stress in English show that, by compressing the unstressed syllable more than the stressed syllable, the most informative part of the word is retained. On a higher level, one could imagine that the same speaker 'trategy' applies. In line with the fast speech timing of lexical stress, words without pitch accent might be compressed more, relatively, than words with pitch accent. If this is the case, the durational correlate of pitch accent also becomes more prominent at faster speech rates. So, in addition, we studied the effect of increasing speech rate on vowels in accented and unaccented condition.

The following hypotheses were tested:

- As speech rate increases, the relative shortening of unstressed syllables is greater than that of stressed syllables. In particular, unstressed vowels will be shortened more than stressed vowels, relatively.
- Unstressed syllables containing schwa will be shortened relatively more than unstressed syllables containing 'ull' vowels.
- The durational correlate of pitch accent becomes more prominent at faster speech rates.

2. METHOD

2.1. Test material

Thirty-two bisyllabic nouns were chosen to measure the durations of stressed and unstressed vowels: half of them with schwa, and half of them with 'ull' unstressed vowels. Stress position and vowel length were balanced, if possible. The words were embedded in long sentences because it is easier to attain a
high speech rate in longer sentences. As speech rate normally decreases towards the end of the sentence, the target words appeared at the beginning of the sentence. The sentences were recorded in two conditions: one in which the bisyllabic target word had a pitch accent, and one in which the word was unaccented. A context sentence preceded the test sentence to indicate which words were to receive pitch accent in the following sentence. The two accent versions of the test sentence are exemplified in (1a-b). Capital letters indicate which words receive pitch accent, and the target word is in bold.
After that, the father swept the leaves in a corner and he could finally sit down on a bench.

In the [-pitch accent] condition, the first adverb of the test sentence received pitch accent. The sentence structure was always the same, and so was the position of the target word.

2.2. Speakers

Four female native speakers of Dutch were asked to read the test material aloud at normal and very fast speech rate. They were paid €20 for their participation.

First, the speakers were asked to read the material at a normal rate. During the reading session the experimenters judged the speaker’s performance. If accentuation was not correct or if the sentence was not read fluently, the speaker was asked to read those sentences once again. After all the material was recorded at the normal rate, the speaker was instructed how to obtain a fast speech rate without abnormal slurring. Speakers were then instructed how to use the stopwatch, so they could get an impression of how fast they could speak, and they could try to excel themselves in their speech rate. In order to increase the speech rate, they were asked to read out each sentence four or five times, and to keep an eye on the articulation time for each trial (i.e. for each combination of context sentence and test sentence). Again, the experimenters judged the speaker’s performance.

The material was recorded onto digital audiotape in a sound-proof cabin with a Sennheiser ME 10 microphone. The signal was filtered with a highpass filter (cut-off frequency of 75 Hz). The signal was then downsampled to 16 kHz.

2.3. Duration measurements

One trial was chosen of each accent version of the test sentence: this choice was based on the correctness of the accent pattern. As there were about four versions of each sentence pair in the fast rate condition, the fastest trial with a correct accent pattern was chosen. The durations of the stressed and unstressed vowels were measured. The vowel onset corresponded with the first positive zero crossing of the first periodic waveform at which an increased amplitude and a change in the wave form occurred. The offset of the vowel corresponded with the last positive zero crossing before the following plosive or fricative started. Segmentation was rather difficult in the fast speech because of heavy coarticulation. In case no periodic vowel signal could be traced in the waveform and spectrogram, a minimum duration of 5 ms was assumed. This was considered a minimum duration because it corresponded to about one period (because the speakers were female with a F0 of about 200 Hz).

3. RESULTS

First of all, articulation rates were computed for the normal and fast speech rates by measuring the duration of the first part of the test sentence (containing the target word) and dividing it by the number of syllables. The average articulation rate was 6.7 syllables/s at normal speech rate, and it was increased to 10.5 syllables/s at the fast rate.

3.1. Speech rate and lexical stress

Our first hypothesis was that as speech rate increases, the relative shortening of unstressed syllables is greater than that of stressed syllables. As we will concentrate mainly on vowel durations, we have checked first whether increasing speech rate had a similar non-linear effect on the level of the entire syllable, and not only on the vowel durations. We measured the syllable durations of the bisyllabic target words in the [+pitch accent] conditions of one speaker. In the fast rate condition, the stressed syllable was reduced to 64% of its normal rate duration, and the unstressed syllable was reduced to 45%. These data suggest that it is indeed the case that the entire syllable is reduced according to its stress level.

Turning now to vowel durations, the mean durations of stressed and unstressed vowels at normal and fast speech rate are shown in Table 1, together with the fast/normal ratios (vowel duration at fast rate / duration at normal rate).

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Normal rate (ms)</th>
<th>Fast rate (ms)</th>
<th>Fast/Normal ratio (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stressed vowel</td>
<td>114</td>
<td>75</td>
<td>0.67</td>
</tr>
<tr>
<td>Unstressed vowel</td>
<td>55</td>
<td>21</td>
<td>0.42</td>
</tr>
</tbody>
</table>

Table 1: Mean durations stressed and unstressed vowel at normal and fast speech rate. Fast/normal ratios are also given.

At fast speech rate, unstressed vowels were reduced to 42% of their normal rate duration, and stressed vowels were reduced to about 67% of their normal rate duration. Hence, we see the similarity with non-uniform effect of increasing speech rate on the level of the entire syllable.

The fast/normal ratios (within each item, per vowel) were analysed in a Repeated Measures ANOVA on the thirty-two items and a Repeated Measures ANOVA on the four speakers, with Stress (Stressed vs. Unstressed) and Accent (Accented vs. Unaccented) as fixed factors. The analyses showed a significant effect of Stress on the fast/normal ratios ($F_1(1,3)=158.6, p=0.001; F_2(1,31)=64.0, p<0.001$).
Half of the bisyllabic words contained twoull’vowels, and these were balanced for vowel length (cf. Section 2.1.). For this subset of items, the fast/normal ratios of the stressed and unstressed syllables showed the same difference: 0.66 and 0.42 for the stressed and unstressed vowels, respectively.

This means that our first hypothesis is confirmed: regardless of vowel length, the unstressed vowel is reduced relatively more in duration than the stressed vowel when speech rate is increased.
3.2. Schwa vs. ull’ unstressed vowels

Our second hypothesis was that schwa is shortened relatively more than unstressed ull’ vowels. In Table 2 the results are shown for the two types of unstressed vowel.

<table>
<thead>
<tr>
<th></th>
<th>duration normal rate (ms)</th>
<th>duration fast rate (ms)</th>
<th>fast/normal ratio (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>full unstressed</td>
<td>66</td>
<td>24</td>
<td>0.42</td>
</tr>
<tr>
<td>schwa</td>
<td>44</td>
<td>17</td>
<td>0.41</td>
</tr>
</tbody>
</table>

Table 2: Mean vowel duration unstressed vowel at normal and fast speech rate (plus fast/normal ratio).

Table 2 shows that the fast/normal ratio of unstressed ull’ vowels equals that of unstressed schwa vowels. The fast/normal ratios of the unstressed vowels were entered into Repeated Measures ANOVAs on items and on speakers, with Vowel Type and Accent as fixed factors. In the item analysis, the items were nested under Vowel Type. Vowel Type (Schwa vs. ull’ vowel) did not have a significant effect on the fast/normal ratios ($F_{(1,3)}=7.07$, $p=0.076$; $F_{(1,3)}=5.62$, $p=0.024$). However, figure 1 shows that there was a trend for [+accent] vowels to be reduced relatively more than the [-accent] vowels, which is exactly the opposite of what we had expected. Although Figure 1 shows a tendency for the unstressed vowels to be affected more by the factor Accent, the interaction between Stress and Accent was not significant ($F_{(1,3)}=30.5$, $p=0.012$; $F_{(1,3)}<1$, n.s.).

The analyses of variance with Stressed vs. Unstressed and Accented vs. Unaccented as fixed factors (mentioned already in 3.1.) showed that there was no significant main effect of Accent ($F_{(1,3)}=7.07$, $p=0.076$; $F_{(1,3)}=5.62$, $p=0.024$). However, in the analyses on the fast/normal ratios of the unstressed vowels only, with Accent and Vowel Type as fixed factors (mentioned in 3.2.), the effect of Accent was not significant either ($F_{(1,3)}=10.0$, $p=0.051$; $F_{(1,3)}=2.71$, n.s.).

So, whereas vowels in words with a pitch accent are longer than vowels in unaccented words at normal rate, this duration difference is almost absent at fast speech rate. This means that although the durational correlate of lexical stress becomes more prominent at faster speech rates, the durational correlate of sentence or pitch accent disappears with increasing speech rate.

3.3. Speech rate and pitch accent

The third hypothesis was that vowels in words bearing a pitch accent would reduce relatively less, with increasing speech rate, than vowels in words without a pitch accent. In Figure 1 the mean vowel durations of the stressed and unstressed vowels are shown, at both speech rates, and in [+pitch accent] and [-pitch accent] condition.

![Figure 1](image)

Figure 1: Mean vowel durations (in ms) of stressed ([+stress]) and unstressed vowel ([-stress]) in both [+pitch accent] and [-pitch accent] condition, at both speech rates. Fast/normal ratios are given above the paired bars.

This duration study was set up to investigate the difference in word-level and phrase-level timing between normal and fast rate speech. Literature on English had shown that speakers tend to reduce the duration of lexically unstressed syllables (and vowels) more than that of lexically stressed syllables and vowels. Our aim was to find out whether this is also true for Dutch.

We expected that the durational correlate of lexical stress would indeed become more prominent with an increase in speech rate, although different results were found for Dutch by Den Os (1988). In her study, increasing speech rate had a greater effect on stressed vowels than on unstressed vowels. However, Den Os (1988) may have underestimated the relative shortening of unstressed vowels because the fast unstressed vowels which were too short to be measured were disregarded. Our study on Dutch showed that, as in English, increasing speech rate affected the unstressed vowels more, relatively, than the stressed vowels.

Nevertheless, we expected a difference between English and Dutch because unstressed vowels in Dutch can be ull’ vowels, whereas unstressed vowels in English are almost always reduced to schwa. We expected that schwa would be more compressible than unstressed full vowels. However, we found that the relative shortening of unstressed schwa was equal to that of unstressed ull’ vowels. Thus, the relative amount to which a vowel is reduced with increasing tempo depends mainly on the stress level of the syllable, and not on the quality of the vowel. As our test material was not suitable to systematically compare the shortening of long vs. short vowels, we cannot answer the question whether long vowels and short vowels are reduced to the same degree.

4. DISCUSSION AND CONCLUSION

This duration study was set up to investigate the difference in word-level and phrase-level timing between normal and fast rate speech. Literature on English had shown that speakers tend to reduce the duration of lexically unstressed syllables (and vowels) more than that of lexically stressed syllables and vowels. Our aim was to find out whether this is also true for Dutch.

We expected that the durational correlate of lexical stress would indeed become more prominent with an increase in speech rate, although different results were found for Dutch by Den Os (1988). In her study, increasing speech rate had a greater effect on stressed vowels than on unstressed vowels. However, Den Os (1988) may have underestimated the relative shortening of unstressed vowels because the fast unstressed vowels which were too short to be measured were disregarded. Our study on Dutch showed that, as in English, increasing speech rate affected the unstressed vowels more, relatively, than the stressed vowels.

Nevertheless, we expected a difference between English and Dutch because unstressed vowels in Dutch can be ull’ vowels, whereas unstressed vowels in English are almost always reduced to schwa. We expected that schwa would be more compressible than unstressed full vowels. However, we found that the relative shortening of unstressed schwa was equal to that of unstressed ull’ vowels. Thus, the relative amount to which a vowel is reduced with increasing tempo depends mainly on the stress level of the syllable, and not on the quality of the vowel. As our test material was not suitable to systematically compare the shortening of long vs. short vowels, we cannot answer the question whether long vowels and short vowels are reduced to the same degree.
In accordance with the non-uniform compression of stressed vs. unstressed syllables, we expected that speakers would shorten unaccented words relatively more than accented words in order to spare the most informative parts of the sentence. We therefore measured vowel durations of bisyllabic words in [+pitch accent] and [-pitch accent] condition. Contrary to our hypothesis, accented vowels were shortened more than unaccented vowels, thereby eliminating the durational correlate of pitch accent at fast speech rate.

One might argue that the vowel durations in [-pitch accent] condition had reached some sort of minimum duration, so that they could not be compressed any more. This could then explain why the accented vowels, which were longer than unaccented vowels at normal rate, could reduce more than unaccented vowels.

An alternative explanation might be that duration is an important cue in indicating lexical stress, but the durational correlate of pitch accent is sacrificed when the speaker is pressed for time. In order to indicate which words are accented in a sentence, the pitch excursion itself is probably a much more important cue than duration (Sluijter 1995).

With respect to the durational correlate of accent, we looked at the vowel durations of bisyllabic content words (nouns) in [+pitch accent] and [-pitch accent] condition. However, if we had compared the reduction of these content words with the reduction of, for example, articles and prepositions in the same phrase, we might have found that the latter are subjected to extra heavy reduction. This could then mean that increasing speech rate is accompanied by a change in timing at the phrase-level, but this change in timing does not involve the durational correlate of pitch accent.

We conclude that increasing speech rate is accompanied by important changes in word-level and phrase-level timing in Dutch. Taking into account some of the non-linearities in the speakers’ way of increasing speech rate might yield valuable improvements for the intelligibility of time-compressed speech (Covell, Withgott & Slaney 1998, and Janse 2000).

5. REFERENCES


Correspondence address: Esther Janse, Utrecht Institute of Linguistics OTS, Utrecht University, Trans 10, 3512 JK Utrecht, The Netherlands. E-mail: Esther.Janse@let.uu.nl