Toneme realization in two North Norwegian dialects

van Dommelen, W. A.

journal: Proceedings of Fonetik, TMH-QPSR
volume: 44
number: 1
year: 2002
pages: 021-024

http://www.speech.kth.se/qpsr
Toneme realization in two North Norwegian dialects

Wim A. van Dommelen1 & Randi A. Nilsen2
1Department of Linguistics, NTNU, Trondheim
2Department of Scandinavian Studies and Comparative Literature, NTNU, Trondheim

Abstract

The goal of this study is to shed some light on the realization of toneme 1 vs. toneme 2 words in two North Norwegian dialects. Two speakers were recorded pronouncing a number of tonal minimal pairs both in a text and as isolated words. In a listening test, a group of listeners identified the words having only prosodic cues at their disposal. For the different words, identification rates varied between chance and almost perfect identification.

Following the listening test, some words judged by the listeners were acoustically analyzed. Words with high correct identification rates showed clearly distinctive F0 contours, in contrast to words that had yielded low correct identification. The results suggest that vowel duration does not play any role in the toneme distinction.

Introduction

Norwegian is a language which to a certain degree exploits tonal differences to distinguish between lexical entities. An often-cited example is the minimal pair bønder (l/bøner/; "peasants") vs. bønner (l/bøner/; "beans"). The realization of toneme 1 and 2 words varies across the different dialect regions and parts of northern Norway are considered to lack tonal contrast (e.g., Almberg, 2001; Fintoft 1987; Jensen, 1961; Kristoffersen, 2000, p. 233ff).

The goal of the present investigation was to shed a preliminary light on both production and perception of tonal accents in two North Norwegian dialects from the Nordland region. On the production side, the question was whether speakers systematically differentiate toneme 1 and toneme 2 and, if so, by what means? As to perception, we wanted to investigate whether listeners are able to consistently identify minimal tonal pairs, i.e. without having any other cues as to lexical identity at their disposal.

Experimental procedure

Preparation of stimuli

Speech recordings were made in the studio of the Linguistics Department using a Milab LSR 1000 microphone and a Fostex D-10 digital recorder with a sampling rate of 44.1 kHz. Manipulations were carried out using the Praat program (Boersma and Weenink, 2001).

The speech material consisted of two parts. Part one was a one-page text written by one of the authors (RAN). The text had the character of a novel fragment and used orthography close to typical North Norwegian pronunciation. Contained in the text were nine different toneme 1 words and their toneme 2 counterparts. Two of the toneme 1 words occurred three times; two toneme 2 words two times each. This yielded a total of 18+6= 24 stimulus words contained in the text.

The second part of the speech material consisted of isolated words elicited in a dialogue with RAN. Here, the speakers were asked to respond to questions like "What is the definite singular form of [stimulus word]?” In total, nine pairs of toneme 1 and 2 words were recorded. Eight pairs were identical with the ones occurring in the text, one pair being substituted with another one. Since one toneme 1 word occurred twice, the total number of words was 2x9 +1= 19.

Using this speech material, two North Norwegian speakers were recorded: a 24-year-old female (B) from Hemnes (Nordland) and a male (H) from Rana (Nordland) aged 24. Both speakers are students at NTNU’s Department of Scandinavian Studies and Comparative Literature. Whereas the male speaker considered...
himself to make a phonemic distinction between toneme 1 and toneme 2 words in his dialect, the female speaker thought she did not make any consistent tonal differences. Specifically, she considered herself to be unable to distinguish between toneme 1 and toneme 2 words in the absence of context indicating the words' lexical identity.

**Listening test**

From the recorded speech material, for each of the two speakers the 43 (24+19) stimulus words were excised for use in the listening test. For both speakers, one toneme 1 word occurring in the text had to be discarded due to strong coarticulatory effects. Since for technical reasons one isolated word produced by the female speaker had to be rejected, the basic material for the test collected from the two speakers comprised 23x2 + (18+19)= 83 words. In the listening test, these 83 stimuli occurred twice in a completely randomized order.

Using the CSRE (Computerized Speech Research Environment) program, the 2x83=166 stimuli were presented to individual listeners seated in the studio of the Department of Linguistics. Their task was to categorize the stimuli as A or B using short descriptions given in a list. In English translation, an example of such a description for the stimulus word *fjæra* is: A - "definite form singular of substantive *a feather*"; B - "definite form of substantive *a low tide*".

The group of subjects comprised 9 listeners with a North Norwegian background and 4 listeners speaking other dialects. They were all familiar with technical linguistic terms as used in the word descriptions. Before performing the test they were given a written instruction supplemented by an oral explanation.

**Results**

**Perception**

In the analysis of the perception test results, the following four factors were taken into account: context (isolated words vs. words excised from context), speaker (B vs. H), toneme (1 vs. 2), and listener dialect (North Norwegian vs. other). The results of an analysis of variance with these factors are displayed in Table 1. In congruence with our expectations, significantly higher identification rates were found for the words in isolation than for the words excised from context. As can be seen from Table 2, the effect of context amounted to 76.3-63.8= 12.5% for the group of North Norwegian listeners and 65.9-57.9= 8.0% for the listeners with other dialectal backgrounds.

<table>
<thead>
<tr>
<th>Factor</th>
<th>F</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context</td>
<td>16.888</td>
<td>1, 1063</td>
<td>&lt;.000</td>
</tr>
<tr>
<td>Speaker</td>
<td>1.611</td>
<td>1, 1063</td>
<td>.205</td>
</tr>
<tr>
<td>Toneme</td>
<td>20.368</td>
<td>1, 1063</td>
<td>&lt;.000</td>
</tr>
<tr>
<td>Listener dialect</td>
<td>10.540</td>
<td>1, 1063</td>
<td>.001</td>
</tr>
<tr>
<td>Speaker x toneme</td>
<td>74.288</td>
<td>1, 1063</td>
<td>&lt;.000</td>
</tr>
</tbody>
</table>

Averaged across all conditions, the factor speaker did not affect recognition rates (mean values of 65.0% and 68.8% for speaker B and H, respectively; F(1, 1063)= 1.611; p= .205). Inspection of the values separated for the two speakers, however, showed that in general toneme 2 words spoken by speaker B were identified better than toneme 1 words. For speaker H the opposite was true. This resulted in a significant speaker x toneme interaction (cf. Table 1; all other interactions turned out to be statistically non-significant). Since for the speaker H tokens the toneme effect was considerably larger than for the speaker B ones, generally higher recognition rates for toneme 1 words were found (toneme 1: 73.2% vs. toneme 2: 59.9%).

<table>
<thead>
<tr>
<th></th>
<th>Isolated</th>
<th>Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>70.0</td>
<td>92.8</td>
</tr>
<tr>
<td></td>
<td>60.2</td>
<td>86.1</td>
</tr>
<tr>
<td>H</td>
<td>79.9</td>
<td>61.7</td>
</tr>
<tr>
<td></td>
<td>65.2</td>
<td>41.9</td>
</tr>
<tr>
<td>Overall</td>
<td>76.3</td>
<td>63.8</td>
</tr>
</tbody>
</table>

(b) Listeners with other dialects

<table>
<thead>
<tr>
<th></th>
<th>Isolated</th>
<th>Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>60.0</td>
<td>82.5</td>
</tr>
<tr>
<td></td>
<td>45.8</td>
<td>72.9</td>
</tr>
<tr>
<td>H</td>
<td>62.5</td>
<td>56.9</td>
</tr>
<tr>
<td></td>
<td>69.3</td>
<td>43.2</td>
</tr>
<tr>
<td>Overall</td>
<td>65.9</td>
<td>57.9</td>
</tr>
</tbody>
</table>
Finally, the group of listeners with a North Norwegian dialect background showed significantly higher identification rates compared with the listeners coming from other dialect areas. The size of the overall effect of listener dialect was 7.9%.

Production

F0 contours

As has been shown in the presentation of the perception data in the preceding section, the identification rates in the listening test varied considerably. In addition to the factors discussed above, large identification differences were found between the various words presented in the listening test. We can take advantage of this variation and look into a possible relationship between identification rates and articulation of particular words. In the following analysis, word pairs containing other than tonal cues were not taken into account. For example, apart from a tonal difference the toneme 2 word anna ("[something] different") was distinguished from its toneme 1 counterpart by palatalization of the nasal. This type of cue naturally helped the listeners in their identification task, but it is not under scrutiny in this study.

The highest identification rates were found for rana (toneme 1 realization: Rana ["name of a county"]; toneme 2: "robbed" [past participle]) with 90.4% correct for the whole group of listeners (North Norwegian listeners: 95.8%; others: 78.1%). Second best was the word pair rosen (toneme 1: "the praise"; toneme 2: "the rose") with an overall rate of 76.4% (North Norwegian listeners: 78.5%; others: 71.9%). Analysis of the first syllable’s vowels showed that their f0 contours were similar, without any obvious differences between the two speakers or the context conditions (isolated word vs. in context). Therefore, for each toneme condition f0 contours were averaged across six vowels (rana isolated and rosen isolated/in context for two speakers). For the sake of comparison, the two resulting mean f0 contours were frequency adjusted so as to achieve the same mean frequency (Figure 1). It appears from the figure that the two toneme contours are clearly distinct. While the toneme 1 contour is characterized by a relatively early top in the vowel followed by a clear fall, the toneme 2 contour rises less steep and reaches its top only at the end of the vowel. The f0 contours in the second syllable of rana (not shown in the figure) were in congruence with this picture, showing a marked falling for toneme 1 and a continuously high f0 contour for the toneme 2 condition.

A typical example of a word pair for which low identification scores were found was fjæra (toneme 1: "the feather"; toneme 2: "the low tide") with an overall identification rate of 59.3% correct (North Norwegian listeners: 66.7%; others: 42.7%).

The fjæra toneme 1 word occurred in total four times, its toneme 2 counterpart two times for each speaker. Figure 2 depicts the mean f0 contours averaged across speakers and conditions. It is obvious from the figure that the two speakers did not make any real tonal distinction between the word pair members. The differences between the two contours are very small so that they safely can be said to coincide.

Figure 1. Mean f0 contours for the first syllable's vowels in rana and rosen (averaged across two speakers; n = 6). For comparison, the contours were frequency adjusted.

Figure 2. Mean f0 contours for the first syllable's vowels in fjæra (averaged across two speakers; n = 8 and n = 4 for toneme 1 and 2, resp.). For comparison, the contours were frequency adjusted.
Vowel duration
The results presented in the previous section suggested that in the two dialects investigated here pitch contour can to a certain extent be used as an auditory cue for the distinction of toneme 1 vs. 2 words. In addition, other acoustic cues might play a role, one obvious candidate being temporal organization. Therefore, as a preliminary step the duration of the first syllable's vowel in the words dealt with above was measured.

For the words *rana* and *rosen* no consistent pattern emerged. Mean vowel durations for toneme 1 vs. 2 words were 128.3 ms and 124.2 ms, respectively. This indicates that the listeners could not rely on vowel duration in their word identification task. In all probability, the main auditory cue for the toneme distinction is the pitch contour.

Contrary to these results, for the toneme 1 vs. 2 realizations of *fjæra* vowel durations appeared to be different: 153.1 ms and 116.2 ms, respectively. However, in spite of this potential cue listeners have not been able to reliably identify toneme 1 vs. 2 *fjæra* realizations. This leads to the conclusion that vowel duration does not play any role in the toneme distinction.

Discussion and conclusions
Though the size of the present study in terms of number of speakers, speech material and listeners is rather limited, some preliminary conclusions concerning North Norwegian toneme realization can be drawn.

The highest correct identification rates found for some words in the listening test (90-95%) possibly represent the maximum that can be expected given the test design: First, the listeners had to identify the stimulus words via the metalevel of linguistic descriptions. Second, non-systematic context effects affecting the prosodics of the words excised from the text rendered identification more difficult. It remains to be investigated to what extent the present test design affected identification performance. This can be done by applying the same design to dialects that traditionally are considered to fully utilize tonemic contrast. In that case, it seems reasonable to expect generally high identification scores, i.e. on a par with the highest ones found in the present study (90-95%).

In contrast to the highly identifiable minimal pairs, a number of pairs showed rather low identification rates. Since this obviously cannot be explained by the difficulty of the listeners' task, this result suggests a generally low functional load of the toneme distinction in the dialects under scrutiny. This interpretation is supported by the idiosyncratic character of toneme realization, showing for the female speaker high identification rates for toneme 2 as compared to toneme 1 words while for the male speaker the opposite was found.

In general, the picture that emerged from the perception and production data was very consistent. It is especially noteworthy that no significant interactions between the results for the two different listener groups were found. This can be taken as supporting the validity of the data. It can be speculated that the tendencies found here will hold for larger groups of listeners, presumably independent of dialectal background.

Finally, the correspondence between the realization of toneme 1 vs. 2 f0 contours and the varying identification rates pointed to the relevant role of f0 contour for the toneme distinction in the two North Norwegian dialects dealt with here.

References