Parameters of tonal variation in and between three Scandinavian languages

Maria Alm¹, Jardar Eggesbø Abrahamsen², Åsa Abelin³, Egil Albertsen², Jacques Koreman²

¹University of Southern Denmark, Sonderborg
²Norwegian University of Science and Technology, Trondheim
³Gothenburg University, Sweden

mhalm@sdu.dk, jardar.abrahamsen@ntnu.no, asa.abelin@ling.gu.se, egil.albertsen@ntnu.no, jacques.koreman@ntnu.no

Abstract

The three Scandinavian languages Danish, Norwegian and Swedish are closely related languages, but they all have their own distinct “speech melody”. Also language-externally, there are characteristic regional speech melodies.

In this paper, we take the literature as a point of departure for discussing and demonstrating what parameters of prosodic structures must be considered in order to account for the tonal characteristics of these three languages and their varieties. A direct comparison is made difficult by different linguistic traditions. We thus discuss what units could be equivalent and suggest terms and definitions for a crosslinguistic comparison.

Index Terms: intonation, Danish, Norwegian, Swedish

1. Introduction

Although Danish (DK), Norwegian (NO) and Swedish (SE) are closely related languages, their intonation systems partially work in very different ways. In this paper, we present and compare the tonal parameters that lead to both crosslinguistic and intralinguistic variation in intonation. We discuss and suggest terms for a crosslinguistic comparison between the three languages.

In Table 1, we have gathered six parameters from the literature. We also offer a quick comparison with English.

Table 1: Tonal parameters of Danish, Norwegian and Swedish and a short comparison with English.

<table>
<thead>
<tr>
<th>Tonal parameter</th>
<th>DK</th>
<th>NO</th>
<th>SE</th>
<th>ENGL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Accents are signaled by pitch movements</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>2 Accent melody is predetermined</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>3 Use of tone accents</td>
<td>No*</td>
<td>Yes*</td>
<td>Yes*</td>
<td>No</td>
</tr>
<tr>
<td>4 Accent melody varies with prominence level</td>
<td>No</td>
<td>Yes*</td>
<td>Yes*</td>
<td>No</td>
</tr>
<tr>
<td>5 Use of end intonation</td>
<td>No**</td>
<td>Yes**</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>6 Functional use of utterance declination</td>
<td>Yes</td>
<td>No</td>
<td>No*</td>
<td>No</td>
</tr>
</tbody>
</table>

*This parameters is represented in some, but absent in other, regional varieties of the language.

**A local tonal gesture at the utterance end is a possible, but optional, feature in many NO or some DK varieties.

Parameters (1-3), which concern accent melodies and the units they form, are discussed in section 2. Combining accent groups into higher-level domains (parameter 4) is discussed in section 3. Tonal parameters with functions at utterance level (parameter 4-6) are discussed in section 4. The results of the discussions are summarized in section 5.

2. Accent melodies and accent groups

Like in English, in Danish, Norwegian and Swedish pitch movements are used in addition to stress to mark prominent syllables (first parameter in Table 1). In the literature, we find three established levels of prominence, even four if we add the nucleus [1], which has recently been suggested for Swedish [2]. For details on tonal prominence levels, see section 3. Here, we will be concerned with the basic phenomenon of accents.

Table 2: Levels of prominence

<table>
<thead>
<tr>
<th>Prominence levels</th>
<th>Feature</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-tonal prominence</td>
<td>Stress</td>
<td>DK, NO, SE</td>
</tr>
<tr>
<td>Tonal prominence 1</td>
<td>Accent</td>
<td>DK, NO, SE</td>
</tr>
<tr>
<td>Tonal prominence 2</td>
<td>“Focus accent”</td>
<td>NO, SE</td>
</tr>
<tr>
<td>Tonal prominence 3</td>
<td>Nuclear accent</td>
<td>SE</td>
</tr>
</tbody>
</table>

In “intonation languages” like English, one and the same word can be pronounced with different accent melodies depending on the context [1]. In Danish, Norwegian and Swedish, accent melodies are predetermined (second parameter in Table 1). Accent melody is not a speaker’s choice among different possibilities for signaling communicative meanings.

In Danish, the accent melody is part of a larger, tonally defined unit which is called tryggruppe (“stress group”). Also the Norwegian aksentfrase (“accent phrase”) is tonally defined, and accent melody is described with respect to the whole phrase. In contrast, the Swedish prosodiskt ord (“prosodic word”) is primarily a metrically defined unit. Primary accents trigger accent melody, but melody is not definitional of prosodic words. We suggest that a tonally defined concept of accent groups can be used equally well also for Swedish.

1) Definition of accent group (AG)

An accent group is a tonally defined unit consisting of exactly one accent and 0–n succeeding unaccented syllables.
The definition in (1) means that each new accent starts a new accent group which normally stretches up to the next accented syllable, which then starts a new accent group. Unaccented syllables at the beginning of an utterance do not belong to any accent group and do not carry any special melody.

Each Danish regional variety specifies a basic tonal pattern of its own for the whole accent group [3: 183], [4: 340]. Each new accent group then repeats this regional tonal pattern. The melody of the accented syllable is part of the larger accent group pattern, which is the object of interest for tonal description in the Danish literature [4: 345].

In example (2), we have used PRAAT [5] to show the accent groups in the Funen Danish question Have you ever seen a robot? The accented syllables show a rise in tone, the following unaccented syllables then fall in tone. The accent groups are: haveyou [Ever], [SEEN a ro•]. (The word “robot” has the stress on the last syllable in Danish).

Table 3: Suggested terms for the level of accent groups in combination

<table>
<thead>
<tr>
<th>Norwegian</th>
<th>Swedish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intonational phrase</td>
<td>Prosodic phrase</td>
</tr>
<tr>
<td>Several “accent phrases”, the last of which contains a “focus accent” [8], [9]</td>
<td>Several “prosodic words”, one of which contains a “focus accent” [7], or… several “prosodic words”, the first or last of which contains a “big accent” [2]</td>
</tr>
</tbody>
</table>

In contrast, Norwegian and Swedish speakers must choose between two lexically diverse tone accents (NO tonelag; SE ordaccenter) (third parameter in Table 1), called accent 1 and accent 2 in this paper. In example (3), the pronunciation of the accent 1 word skravet (“the document”) and the accent 2 word skrive (“(to) write”), both with the identical segmental representation [ski:və], is demonstrated for a speaker of Sunnmøre Norwegian, which is a high-tone dialect:

Norwegian research traditionally distinguishes between low-tone and high-tone dialects [6], depending on the starting tone in the accent 1 melody. For Swedish, four pairs of regionally varying accent melodies for accent 1 and accent 2 have been suggested, plus some varieties not using any lexically diverse tone accents [7].

Table 4: Cues for focus accents vs. big accents

<table>
<thead>
<tr>
<th>Norwegian</th>
<th>Swedish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-tone dialects like Oslo: increasing pitch scaling of final H in ¹LH and ²LHL</td>
<td>Central SE [2]: melody switch</td>
</tr>
<tr>
<td>Acc1 HL* → big L^H</td>
<td>Acc2 H^L → big H^L</td>
</tr>
<tr>
<td>Sunnmøre [10]: for instance, lowered final L in ¹HL, ²LHL, downstepped H in post-focal AGs that are not inter-focal</td>
<td>Scania [11: 69]: increased pitch scaling of accent melodies</td>
</tr>
<tr>
<td>Ytre Sogn [10]: at times rising the H in ¹HL, ²LHL</td>
<td>¹LH, ²LHL</td>
</tr>
</tbody>
</table>

The association between tone accent type and specific words is in part lexically, in part phonotactically determined. The rules for the distribution of the two accent types over different words vary between regional varieties of the same language (see for example [7: 49-54]). So do the specific accent contours that are associated with accent 1 and accent 2. In all, this means that Norwegian and Swedish speakers do not have any “free choice” of accent melody, either.

3. Accent groups in combination

The next step consists in combining accent groups with each other (the fourth parameter in Table 1).

In Danish, accent groups are lined up one after another, as many accent groups as are needed to finish the message. An utterance can consist of just one accent group, or it can comprise several accent groups. At this level, Grønnum [4] speaks of “phrases” and “utterances” depending on the function of the accent group combination. She uses these two terms for higher-level domains intuitively.

In Norwegian and Swedish, accent groups are combined in at least one additional step before the communicative level of utterance. The next domain can thus comprise more than one accent group. The (best-known) terms used for this higher-level domain and its definitional feature used in Norwegian and Swedish literature are presented in Table 3:

However, since the smaller units in these frameworks, “accent phrases” and “prosodic words”, respectively, do not have comparable definitions, segmenting one and the same utterance into Norwegian intonational phrases vs. Swedish prosodic phrases would yield different analyses. We keep to the tonal definition of accent groups and define the higher-level tonal domain as follows:

4) Definition of a higher-level accent domain

The higher-level accent domain comprises 1–n accent groups as defined in (1), one of which (NO: the last one; SE: first or last one) features a higher-level accent.

The higher-level accent is signaled by different kinds of cues in Swedish and Norwegian varieties, some of which are presented in Table 4. In Central Swedish, even accent melodies change between the levels, which seems rather unique:
In Norwegian, the traditional term “focus accent” for the higher-level accent seems to be appropriate because these accents are assumed to have a direct relation to information-structural focus [11]. However, the appropriateness of the traditional term “focus accent” [7] has been questioned in recent research on Central Swedish [2]. As can be seen in Table 4, there are two melodies for accent 1 and two melodies for accent 2 in Central Swedish. The choice between the two melodies for each accent type is determined by the prominence level; but Myrdal and Riad [2] argue that higher-level prominence in Central Swedish is not automatically associated with information-structural focus. They present examples in which Central Swedish higher-level accents are also used in information-structurally backgrounded parts of an utterance. In (5a-b) below, one of their examples is presented [2: 134 #32a] (this is one of many possible pronunciations of the utterance). In example (5b)...

- the parts of the utterance that form the information-structural background (G) vs. focus (F) are in brackets.
- words with an accent (regardless of prominence level) are marked with ò (for “prosodic word”).
- words with a higher-level accent are marked by small capitals in addition to ò. The higher-level phrases that these accents are heads of are marked with parenthesis and φ (for “prosodic phrase”).

As can be seen in (5b), higher-level accents are used on two words in the backgrounded part of the utterance, bruna (“brown”) and ungar (“babies”); and then also on parken (“the park”) in the focused part of the utterance. That the higher-level accents on bruna and ungar are not some kind of information-structural narrow focus, but serve higher-level chunking only:

5) a. Translation of the utterance in example (5b):
Den bruna haren med många ungar bor i parken.
The brown hare with many babies lives in the park.

b. Prosodic words, prosodic phrases and information structure:
[Den BRUNa haren ø med mänga UNGARøøøø φ]
[bor, ø PARKENøøøø φ]

There is an indirect relation between higher-level accents and information-structural focus in the sense that it takes a higher-level accent to signal focus, but “focus accent” as a general term leads to confusion. Riad and Myrberg instead suggest the theory-neutral terms “small accents” vs. “big accents”. We use “big accent” for Swedish, but in opposition to [2], we define the resulting domain tonally, see (4) above. We call it “big-accent group” because the Swedish term “prosodic phrase” is not compatible with our tonal definition. At the same time, the Norwegian term “intonational phrase” is based on information-structural focus, and in addition, this term also leads to confusion with the utterance-level unit, see below.

The Swedish big-accent group is not yet at utterance level. For example, it lacks final boundary tones [2]. At the utterance level, the last big accent is upscaled to what Riad and Myrberg call a “nuclear accent”, and this is the head of what they call the “intonational phrase”, see next section. In Norwegian literature, there is no discussion of an upscaled nuclear accent. If an utterance comprises more than one big-accent group, the focus accents of these big-accent groups are considered equal [12].

4. The communicative utterance level
For this level, the following terms are used in the literature:
- DK: Utterance (used in an intuitive way) [4]
- NO: Intonational utterance [12]
- SE: Prosective utterance [7], intonational phrase [2]
- Engl/international literature: Intonational phrase [1]

We will adopt the term “intonation-group” preferred by Cruttenden [1]. Although the term “intonational phrase” seems to be the most widely used term, its wording is identical to the term for higher-level accent groups in Norwegian literature (see section 3).

The concept of “intonational group/phrase” seems to be used together with the concept of “nuclear accent” [3, 12]. Nuclear accent is defined as “the pitch accent which stands out as the most prominent in an intonation-group” [1: 42], i.e., we are dealing with prominence levels (fourth parameter in Table 1). Judging from the examples in the literature [1, 2], it can be used both for accents signaling information-structurally broad focus and for accents signaling information-structurally narrow focus. This makes it a useful term for Swedish in comparison to the traditional “focus accent” with its ambiguity between utterance-internal chunking and information-structural focus. However, the concept of nucleus implies that the last big accent of an utterance really is more important than the others and as a metaphor, “nucleus” implies that there is only one nucleus in each intonation-phrase, which then is also suggested in [2].

4.1. Accents and information-structural focus
As demonstrated in section 3, it is possible to use several big accents in an utterance in Swedish. Some of the big accents seem to chunk the speech into big-accent groups without signaling information-structural focus; but one big accent will typically be especially prominent and signal information-structural focus. In [2], this is called the “nuclear accent”. The last, particularly prominent big accent in an utterance is also often called “sentence accent” in utterances with information-structural broad focus; and “contrastive accent” or “emphasis” for a narrow focus (this terminology is for example used in [4]).

In Central Swedish, nuclear accents are upscaled such that the f0 maximum of this accent is higher than in previous big accents, although it has not been experimentally investigated yet whether this is also a perceptive distinction [2: 137]. Post-nuclear accents are always of the “small accent” kind [2: 136]. In Scania Swedish, the focus-signaling accent is slightly upscaled. Post-focal accents are usually downgraded; their pitch range is slightly compressed, and sometimes, they may get de-accentuated and do not carry accent melodies at all [13].

In the Norwegian tradition, all higher-level accents are considered to be information-structurally motivated. “Focus accents” are thus found at places where information structure is projected into syntax in agreement with certain rules that cover both wide focus, narrow focus and double focus [9]. In South-East Norwegian, any non-focal accents after the first focal one downstep the final H tone in each accent melody, creating a phonetically “compressed” pitch contour. If there is more than one focus accent in an utterance (two is the maximum), they are treated as phonologically equal.

In Danish, “sentence accents” are not used [4]. In broad focus, the last accent of an utterance is not upscaled. In narrow
4.2. Final boundary tones

In many languages, intonation-groups are also characterized by having final boundary tones (fifth parameter in Table 1). In (6), we demonstrate the differences in Scandinavia on the question *What education do you have?*

In Swedish, final boundary tones are obligatory [2], [3]. The Central Swedish word *utbildning* ("education") is pronounced with a big accent 2 melody H*LH. The first tone movement H*L takes place on the primary-stressed syllable *UT*. The last high tone H is associated with the secondary-stressed syllable of the word, (UT)-bild-(ning). The low tone at the end is a low boundary tone.

In Norwegian, final-boundary tones are optional and under certain usage restrictions. The low tone at the end of the big-accent group *(UT)danning du he* ("education you have") in (6) is simply the last tone of the Summøre accent 2 melody LHL (see Table 4). It is not an additional final boundary tone.

In Danish, final-boundary tones are not used in most varieties. Note the middle-range tone at the end of the Danish word *UD-dannelse* ("education"). It is the same tone range as in the word *hvad* ("what") at the beginning of the utterance.

4.3. Utterance declination as a functional cue

Declination can be found in all three Scandinavian languages, but in Danish, it is suggested to be used for signaling the communicative function of the utterance [3], [4] (sixth parameter in Table 1). In opposition to final boundary tones, this is a “global” signal of communicative functions [4].

The communicative functions that get signaled in this way concern the difference between statements (steep declination) and different kinds of questions, for example wh-questions (rather steep declination), yes/no questions (moderate declination) and declarative questions (no declination) [4: 348]. Grønnum calls this “intonation” the slope that is formed by the accented syllables after the local tonal movements of each accent group have been “peeled off” [4: 343]. We demonstrate this aspect on the declarative question *So you started school when you were six?* in Funen Danish:

5. Summary

In this paper, we have discussed terms and definitions for six tonal aspects that lead to differences in speech melody between and within Danish, Norwegian and Swedish.

Pitch movements are used on stressed syllables to signal prominence. This is called accent. Accent melodies are predetermined by the regional varieties. The domain formed by the tonal influence of an accent melody we call the accent group. In Danish, rather the larger accent group pattern is of interest. Each regional variety has one basic accent group pattern. In Norwegian and Swedish, the difference between the two lexically diverse tone accents always has to be signaled. In Norwegian, the tonal movements of accent melodies can be distributed over the whole accent group, whereas in Swedish they are primarily limited the beginning of an accent group.

In Norwegian and Swedish, accent groups are chunked into big-accent groups, defined by the presence of a higher-level accent. In Swedish, but not in Norwegian and Danish, an upscaled “big accent” signals information-structural focus at utterance level. In Danish, no higher-level accents are assumed.

The unit at utterance level we call the intonation-group. In Swedish, intonation-groups have obligatory final boundary tones. In Norwegian, final boundary tones are optional. The upscaled high or low last tone of the Norwegian often utterance-final focus accent melody seems to make final boundary tones largely superfluous. In Danish, final boundary tones are not used in most varieties. Instead, the declination of an utterance can be used for signaling the communicative function of the utterance as a whole by having different degrees of steepness.

6. End notes

*In Swedish, the prosodic word can comprise syllables both preceding and following the primary-stressed syllable [2], [7].

7. Acknowledgements

This paper is the result of the project *Developing Teaching Material for Speech Prosody to Increase the Mutual Comprehension between Danish, Swedish and Norwegian*, funded by Nordplus.
8. References


[12] The Trondheim Model of Intonation, developed by Thorstein Fretheim and Randi A. Nilsen, see [8: Ch. 10] and [9].