

A Road Map for the Development of Teaching Material on Scandinavian Prosody

Maria Alm¹, Åsa Abelin², Jardar Eggesbø Abrahamsen³, Egil Albertsen³, Jacques Koreman³, Oliver Niebuhr¹, Kerstin Fischer¹

¹University of Southern Denmark

²Gothenburg University

³Norwegian University of Science and Technology

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

ABSTRACT

In two recent Nordplus projects, we have been working on how to teach Scandinavian prosody. One aim was to familiarize young speakers of Danish, Swedish and Norwegian with the speech melodies of their “neighboring languages” since they have increasingly less contact with the other Scandinavian languages and the mutual comprehensibility of the neighboring languages is decreasing. Another, related aim was to find a way to teach prosody to second-language learners of Scandinavian languages. In this paper, we present what we see as a road map to the development of comprehensive teaching materials for Scandinavian prosody: the aspects that need to be addressed and the resources we believe are needed for the development of a teaching material that will appeal to both language teachers and language learners.

Key words: Scandinavian languages, Danish, Swedish, Norwegian, prosody, intonation, speech melody, second-language teaching

1. Introduction

In two recent Nordplus projects, we have been working on how to teach Scandinavian prosody. We had two target groups in mind: on the one hand, young native speakers of Danish, Swedish and Norwegian, on the other hand second-language learners. Both target groups would benefit from getting acquainted with the sound of these languages.

Swedish, Danish and Norwegian are traditionally thought of as “neighboring languages” in their respective countries. The notion of a neighboring language means that speakers of one Scandinavian language can understand speakers of another Scandinavian language without instruction (albeit possibly with some effort). This mutual intelligibility allows Scandinavian speakers to engage in cross-national communication, each speaking in their native language. However, there is growing evidence that mutual cross-Scandinavian comprehensibility is decreasing, especially among young Scandinavians (see for example [1], [2], [3] and [4]). Interest is also decreasing. There seems to be little contact between the language speakers [2, p. 3]: “It becomes clear that young people in Scandinavia in general make little use of the possibilities to have contact with the neighboring languages.” Interest is rather focused on English-speaking cultures. This trend is fueled, for example, by the fact that many of the programs and films shown on TV are broadcast in English (TV programs are not synchronized in Scandinavia),

popular music, also by Scandinavian artists, has English texts, and young people learn English to participate in international computer-gaming communities. Young Scandinavian native speakers tend to speak English with each other [5]. Bohn & Jørgensen [4] find that young Danes have significantly more problems understanding Norwegian and Swedish questions than English questions, and they also need significantly longer to process questions in Norwegian and Swedish than questions posed in English. On the other hand, the authors do not find this finding surprising:

[E]ven though school-age Danes, Swedes, and Norwegians have some obligatory instruction in the other Scandinavian languages, this instruction is minimal, and they do not have much exposure to the other Scandinavian languages in everyday life /.../. In contrast, Scandinavians are instructed in English at school from a very young age (in Denmark now from the 1st grade), which makes it likely that they understand English more effortlessly than they do the neighboring Scandinavian languages. [4, p. 1741]

Examples for the decreasing comprehensibility are ubiquitous. From conversation analytical research, the often observable switch to English in inter-Scandinavian interaction has even been interpreted as an instance of a general principle governing

behavior in communication saying “you should answer in the same language in which you were addressed” [6]. Unfortunately, we think that the language switch between Scandinavian native speakers is due just as much, or more, to the fact that without much previous exposure, it takes a lot of effort to understand the neighboring languages straight-off in oral communication, and to the fact that it is embarrassing and negative for the interpersonal relation to ask “sorry?” too many times. In the article “Why don’t the Danish understand Norwegian?” [7], the dialect researcher Krister Vasshus nails it: “the less you hear and are exposed to a language, the less you feel secure with it” (translation MA).

The mutual comprehensibility between young Scandinavians is so low that Bacquin & Christensen, based on their experience in teaching Danish at Lund University, suggest that you should discard the idea of neighboring languages and instead promote Danish as a “foreign language that is easy to learn” [3].

We think it would be a pity to give up on mutual comprehensibility, which comes almost for free. However, there is clearly a lack of exposure to the neighboring languages. If a teacher does not have a personal interest in the other Scandinavian languages, the topic can be done away with quickly enough in school [8]. Besides, the exercises in the textbooks used in school and upper-secondary school often focus on comparisons of written language. Henriksen reports on another method, though: “The teaching of the neighboring languages is often conducted like native-language teaching: you read an original text (aloud) and subsequently discuss it” ([8, p. 117], translation MA). But texts do not introduce the students to the sound of the neighboring languages. Thus, texts do not help with comprehensibility in oral situations.

We believe that easily accessible, good teaching materials for oral Scandinavian for school and upper-secondary school can help teachers and students to get better acquainted with their neighboring languages. There are some freely accessible oral resources already, for instance the laudable initiative of the digital teaching platform *Norden i skolan* (“the Nordic countries in school”) by Thomas Henriksen, Carl Liungman and Kristín Ólafsdóttir [9]. This platform is accessible in all three languages (<https://nordeniskolen.org/sv/>) and contains both videos and teaching materials for different age groups. However, in our projects we decided to focus on prosody. Intonation is the basic “packaging” of the languages. As Grønnum [10, p. 16] puts it, stress and intonation form the skeleton that holds the individual speech sounds in place. In addition, Grønnum [11] points out that intonation is an immediately obvious and pervasive feature that can be identified also when you cannot make out the words being spoken. A Danish native speaker can recognize the origin of the speaker by his/her regional tonal figures across a room full of people [11, p. 340]. This also means that speech can appear unfamiliar already by its speech melody.

The aims of our two projects were thus to find ways to familiarize young speakers of Scandinavian with the sound of their neighboring languages as well as to support second-language learners of Norwegian, Swedish and Danish. Just like the native speakers, the latter group could benefit from familiarization for reasons comprehensibility and benefit from an easily accessible description of the intonation of their target language. Thus, second-language teachers and second-language learners were also regarded as target groups. In one of the

projects, we developed a freely available, online tool, which you can speak directly into and immediately get a visualization of the intonation and of the placement of the stresses of that utterance. In the other project, we examined the tonal features of the three languages and developed a sample of a freely available teaching material for universities, upper-secondary schools and language classes, or it can be used for self studies.

In this paper, we present our experience with what is needed for developing more extensive teaching materials for Scandinavian prosody that will appeal to, and be useful for, both language teachers and language learners.

2. The Road Map in Practical Application

In our work towards this road map to teaching Scandinavian prosody, we found it necessary to address several aspects before teaching materials addressing intonation explicitly can be developed in the first place: Awareness raising, so that learners learn to hear the relevant distinctions; teaching the specific features in which the three languages differ; deciding on what varieties to teach; building a visualization tool that helps students understand what they are producing and providing online feedback on their productions.

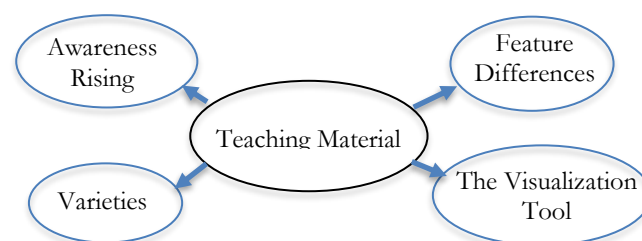


Figure 1: *The topics in the teaching material.*

Awareness raising is necessary because many non-linguists have no idea what intonation and prosody are. Our usability tests of the tool [12] we developed with language teachers revealed that the term “speech melody” is hardly more informative to many, and even longer explanations leave them confused. Thus, teaching material for prosodic information should allow learners to start with listening to examples rather than introducing theory. The material we developed starts with practical examples demonstrating what stress is and then move on to comparing the same utterances in all three languages, with focus on stress and intonation. In this part, we have included a small sample of common, short phrases, like “Hi!”, “That’s great!” and “Are you crazy?” which all students can try out for themselves. The examples are numbered, but the teacher or learner should be able to choose freely where (s)he wants to start and jump back and forth between the topics.

The creation of teaching materials furthermore necessitated that we understand in what ways the three languages differ from each other and develop a common language to talk about them, for example, about accents and accent groups. Swedish and Norwegian lexically diverse tone accents are also explained here. These are the topics explained in section 3 below.

In the creation of teaching material, the researcher also needs to decide what varieties to teach. There are, for instance, attitudinal differences towards different varieties, which may influence the choice of varieties presented, as well as

functionality and distribution, among other factors. We discuss this in section 4.

A third factor concerns the fact that not all tonal characteristics of Norwegian, Danish and Swedish may influence comprehension to different degrees and may thus be relevant to different extents. This is discussed in section 5.

A fourth factor we considered was how to present the relevant tonal features to language learners in a comprehensible way. In section 6, we present the visualization tool we developed based on these considerations.

In section 7, we briefly introduce our sample for a teaching material, and in section 8 we summarize this paper and present some additional thoughts on what is needed for a functioning introduction of the spoken neighboring countries in the school system and in second-language classes.

3. Identifying the Relevant Tonal Features and Defining Cross-Linguistic Terminology

In order to develop teaching materials for the prosody of the Scandinavian languages, it needs to be determined what the characteristic features are in which the languages differ. Danish, Norwegian and Swedish share important basic features, but there are also cross-linguistic and intra-linguistic differences. In addition, separate linguistic traditions for prosodic description make a direct comparison between the three languages difficult. Different traditions and schools focus on different issues, and both terminology and definitions vary between the traditions.

Based on the literature, we have identified six tonal features of importance for describing the intonation of the three languages. These tonal features were presented and discussed in detail in [13], in which we also made a first suggestion for a terminology for cross-linguistic comparison. Table 1 is reduplicated from [13, p. 1] and comprises also a comparison with English, a language that many people have learnt to speak as a foreign language.

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

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

*This parameter 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 Norwegian and some Danish varieties.

All three languages make use of stress (signaled by several different features such as duration, quality and intensity) to form rhythmic units. In common approaches, a distinction is made

between primary-stressed, secondary-stressed and unstressed syllables. All three languages, and English, also make use of tonal gestures on primary-stressed syllables. This is called accent, see tonal parameter 1 in Table 1.

In English, the accent melody of a primary-stressed syllable can be varied to convey subtle communicative meanings [14]. Not so in the Scandinavian languages. Here, the tonal gestures that are triggered by primary-stressed syllables are predetermined, see parameter 2 in Table 1.

In Danish, there is one basic, very characteristic tonal pattern for each regional variety (parameter 2). The speaker usually starts that one basic pattern anew with each new primary-stressed syllable that comes up in an utterance [11, p. 340]. The tonal pattern is described for the whole “accent group” [11], [10], which starts with a characteristic tonal gesture on a primary-stressed syllable and then comprises all the following non-primary-stressed syllables, until the next primary-stressed syllable comes along and initiates a new accent group.

Also in most Norwegian and Swedish varieties, accent melodies are predetermined (parameter 2), but here the situation is more complicated. In these languages, lexically diverse tone accents are used, see parameter 3 in Table 1. This means that accent melodies come in pairs and stand in opposition to each other. The use of tone accent 1 or tone accent 2 in different words is determined by lexical specifications as well as by morphology and phonotactics, and not by the individual speaker. In some cases, the choice of tone accent can lead to minimal pairs. In Figure 2, two different grammatical forms of the adjective *fin* ‘beautiful’ are demonstrated in Southern Sunnmøre Norwegian. In this variety, the words are segmentally identical, but with tone accent 1, *fine* [‘fi:nə] means ‘beautiful-singular’; and with tone accent 2, it means ‘beautiful-plural’:

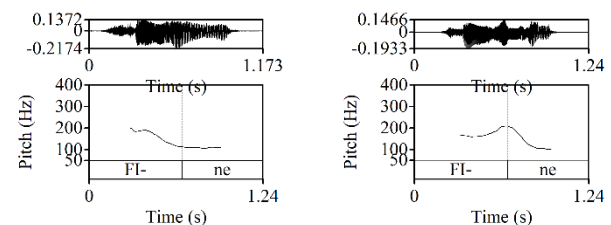


Figure 2: *The minimal pair fine (accent 1, ‘beautiful-SING’) and fine (accent 2, ‘beautiful-PLUR’) spoken in Southern Sunnmøre Norwegian.*

In all three languages, accent melodies are triggered on primary-stressed syllables. For non-native listeners, this the primary stress in combination with the launch of an accent melody make for a practical starting point for the description/analysis of Scandinavian intonation. We therefore adopt the Danish [10, Ch. 8], [11, Ch. 18] and Norwegian ([15], [16], [17], [18, Ch. 10]) tradition of basing the accent group on tonal criteria: on a primary-stressed syllable, the tonal pattern changes, and a characteristic accent (group) melody is used on a group of syllables until the next accented primary-stressed syllable occurs.

Stress is a phenomenon with different levels of prominence (primary and secondary stress), and in Scandinavian languages, accents are used to signal primary stress. Also, in both Norwegian and Swedish, details of the accent melodies can be varied to signal what can be informally labelled “very important

information”, i.e. an even higher prominence level, see parameter 4. The accent of higher-level prominence is traditionally called “focal accent”. However, “focus” today is often understood as having the function of signalling information-structural focus. This is compatible with the Norwegian use of “focus accent”, because in Norwegian all uses of the higher-level accent are considered to be associated with information-structural focus [19]. (Please note that “focus” should not be confused with “rheme” (i.e. information that the speaker assumes is new to the hearer), see [20]). In contrast, in Swedish the higher-level accent can be used also in parts of an utterance that are clearly backgrounded from the information-structural perspective (cf. [21] and [13]). Finding crosslinguistic terms for the higher-level accent realization and its domain has proven difficult, and in [13] we ended up using different terms for Swedish and Norwegian.

Higher-level accents are signaled by different kinds of cues in Swedish and Norwegian varieties, for example by increasing pitch scaling in different ways, see [13]. In Central Swedish, there are even different melodies between the two prominence levels. This makes for two pairs of opposing accent melodies: accent 1 and accent 2 melody on the lower accent level; and accent 1 and accent 2 melody on the higher accent level [21].

In Figure 3, a Central Swedish utterance has been visualized using PRAAT [22]. The utterance is translated in example (1) below Figure 3. The first text tier shows how the syllables are aligned with the speech signal. Accented syllables are marked with capital letters: *RO-bo-tar*. The second tier shows the tonal target points of the accent melodies. The third tier shows the extension of the single accent groups. The fourth tier shows the higher-level accent domain created by the presence of a higher-level accent:

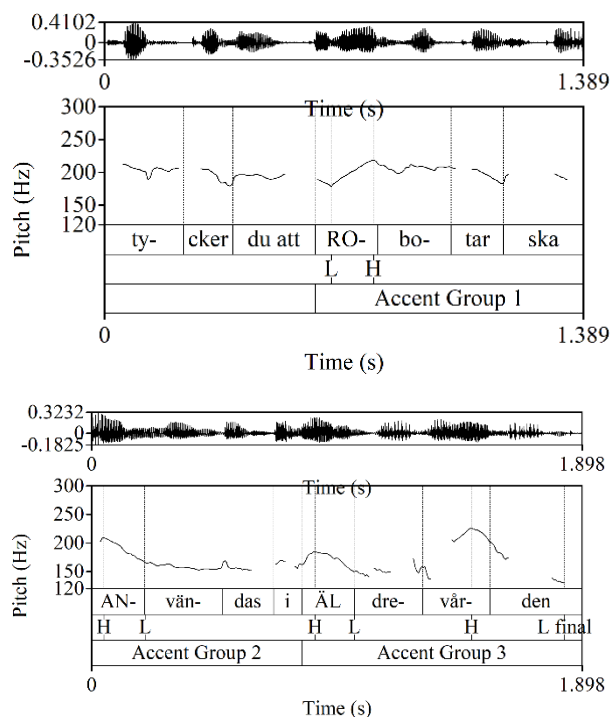


Figure 3: Accents, accent groups and end intonation in an utterance spoken by a female Central Swedish speaker.

In this utterance, there are two higher-level accents and one lower-level accent. There is a higher-level tone accent 1 (LH) on

RObotar and one higher-level tone accent 2 HL+H on the last word *ÄLDrevården*. The HL part of the movement starts on the primary-stressed syllable *ÄL-*, and the last tone H is realized on the last secondary-stressed syllable of the lexical word: *ÄLdre-vård-en*. In between is a lower-level tone accent 2 (HL) on *ANvändas*. On the last syllable of the utterance (*-en*), there is a low final boundary tone. (Creaky voice is used on several places in the speech signal, and this is seen as discontinuities and disturbances in the visualization.)

In example (1) below, the utterance in Figure 3 is translated, using the following transcription:

- Accented syllables are marked with capital letters: *RObotar*.
- Higher-level accented syllables are additionally marked with bold face: **ÄLDrevården**.
- Single accent groups are marked with square brackets: []

The utterance starts with prelude, a purely rhythmic group *ty-cker du att*. Perceptually, there is a clear secondary stress, i.e. stress without an accent, on the first syllable *ty-*. This prelude is prior to the first accent group, as accent groups start with an accent:

(1) *tycker du att [RObotar ska] [ANvändas i]*
 think you that robots should be-used in
 ‘do you think that robots should be used in’

[**ÄLDrevården**]
 elderly-care
 ‘the care for senior citizens’

Following our definition above, the “accent groups”, which contain precisely one accented syllable (on which either a lower-level accent melody or a higher-level accent melody is realized), makes for a clearly defined domain: they start with an accented syllable and end before the next accented syllable. In contrast, the extension of the higher-level accent groups, a domain which may comprise several (single) accents groups, is much more difficult to define. If there are no information-structurally narrow foci in the sentence, a higher-level accent will most often be found towards the end of the utterance and take scope backwards over the whole utterance. In the Norwegian Trondheim model, the higher-level domain always ends in a “focal” higher-level accent group. Phonologically, this grouping can be seen in declination patterns [18]. However, as mentioned above, in Swedish utterances higher-level accents can also be found in the information-structurally backgrounded parts of the utterance. In Figure 3 above, the LH tonal movement on the accent group *RObotar ska* (“robots should”) is the melody of the Central Swedish higher-level accent 1 [21], and still it is used in an information-structurally backgrounded part of the utterance. The pragmatic focus in the utterance is on the last accent group **ÄLDrevården** (“care for elderly citizens”). The extension of each higher-level accent domain, and the function of the higher-level accents, is then much less clear.

The question is what non-native listeners need to know about accents and higher-level accents for listening comprehension in a real-time communicative situation. This question is left for future research.

In [13], we also discussed the possibility of an utterance-level “intonation group”, as in English [14]. This is suggested for

Norwegian in the Trondheim Model [19: “intonation utterance”], but such a group has not been thoroughly researched yet for Swedish [21].

In Danish, there is no higher-level accent [11], [10]. The last accent of an utterance is of the same kind as previous accents, and narrow focus is not signaled by higher prominence of the focused word, but rather by compressing or completely deaccentuating the next following accent.

In Swedish, and optionally in Norwegian and in some Danish varieties, end intonation is used to signal the end of an utterance, see parameter 5 in Table 1. This means that the tonal movement after the last accent group suddenly “breaks off” and either “reaches into the limits of the speaker’s individual pitch range” [23, p. 296], or the melody turns up and ends in a much higher tone than could be explained by the accent melody itself. This can be seen in Figure 3 above: in the last accent group *Äldrevården* (‘care for senior citizens’), there is a fall to a low, creaky tone on the last syllable.

Unlike Swedish and Norwegian, Danish seems to put the declination throughout an utterance to use as a signal for communicative function [11], [10], see parameter 6 in Table 1. The accented syllables of an utterance will form a downward slope of varying steepness: In statements, the accented syllables form a steep declination, but in questions, the declination is less steep. In yes/no questions without subject inversion, which have the same syntactical format as declaratives, the declination can be almost horizontal. According to Grønnum, the more clearly a question format is marked as a question by syntax and morphology, the steeper the declination is likely to be.

In Figure 4, a male Copenhagen Danish speaker produces two different question types after each other. He starts out with a verbless WH question from the interview script, *How about another dialect?*, which he immediately reformulates into a Yes/No question with subject inversion: *Do you speak another dialect?*

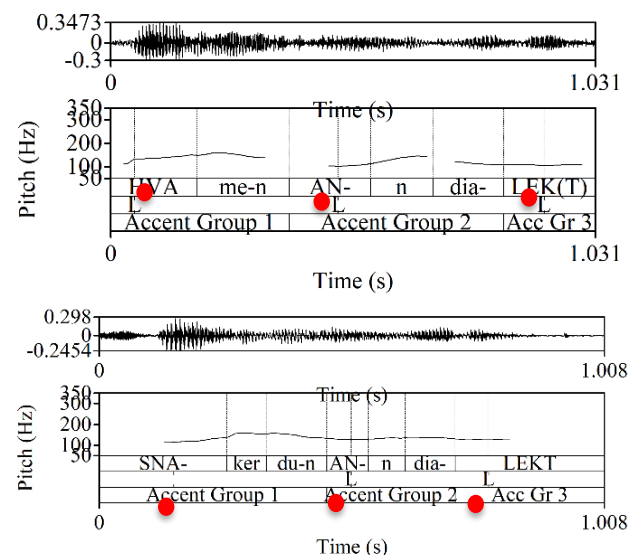


Figure 4: The declination through the stressed syllables in two different question types by the same speaker. First a verbless WH-question, then a Yes/No question with subject inversion. The dots in the figure mark the stressed syllables.

According to Grønnum’s research [11], [10], the accented syllables of WH-questions form a steeper declination than those of Yes/No questions. However, the utterances in our material do not form declinations with clear steepness degrees like those described by Grønnum. In Figure 4, we have measured the lowest tone of the accented syllables because Copenhagen Danish is characterized by having a low tone on the primary-stressed syllables (the tone is measured in the vowel of the accented syllable). The frequency levels of the primary-stressed syllables of the WH question are 135, 103 and 105 Hz. In the Yes/No question, the frequency levels are 116, 128 and 128 Hz:

The fact that declination tendencies are unclear in our speech sample could be due to the fact that our speakers do not read aloud from a list of isolated sentences out of context, but engage in small interviews with each other. This more often brings in other pragmatic factors like narrow focus and emphasis, which can influence how accents are produced. It could also be that the rather small size of our material is too small for clear patterns to emerge.

4. Language-Internal Variation (Varieties)

Pronunciation and intonation do not only vary between languages; they also vary within one and the same language. The question of language-internal variation is connected both to the question tonal characteristics and to the practical considerations behind designing a teaching material: Among all the regional and social varieties of a language, not just the Scandinavian ones, which one(s) should you base a teaching material on?

Due to increasing contact across dialect regions, many Swedish and Danish speakers nowadays speak a vager regional variety rather than dialect [24]. It is characterized by having a nationally standardized vocabulary and grammar. What remains as regional characteristics for spoken language is mainly a varying pronunciation of speech sounds and intonation [25], [26]. In both Denmark and Sweden, people have an idea of a standard variety (Danish *rigsmål* ‘(spoken) language of the country’, Swedish *rikssvenska*, ‘Swedish of the country’ even though there is no official norm, and a standard pronunciation is not taught in school (in contrast to the written standard language). In Denmark, the “standard variety” is Copenhagen Danish, in Sweden it is a blend of the pronunciation used in the larger area around Stockholm. In Norway, the situation is different. Norwegians hold on to their spoken dialects. However, our Norwegian project members report that South-East Norwegian, used among other places in the capital Oslo, seems to have a relatively high status in Norway, if yet subconsciously.

It is therefore not surprising that there is a dominating variety in sound files and pronunciation exercises in existing text books for Swedish and Danish. In Norway, the practical problem of teaching pronunciation is bigger because the variation is richer and there is no clear acceptance of an unofficial standard. The problem is often solved by starting with pronunciation teaching based on the written language standard, which is either *bokmål* or *nynorsk*, depending on the municipality in which the second-language learners live. It varies how dialectal traits are incorporated in the teaching [27].

The arguments in favor of and against using a “standard” variety in teaching materials can be summarized as follows:

Table 2: *Arguments for and against basing teaching materials on a “standard” variety of Danish, Swedish and Norwegian.*

Arguments in favor	Counter-Arguments
There are unofficial “standard” varieties	There is no official pronunciation norm
Variety of the political elite = neutral or high-prestige	We have no right to promote one specific variety
The varieties in the capitals are better researched, described and understood	Keeping to the convenient alternatives will only reinforce the situation
<i>Speaking:</i> You cannot teach all varieties, and teachers cannot speak all varieties	<i>Listening:</i> students/learners are likely to hear different varieties
<i>Speaking:</i> How useful is a low-prestige variety for learners?	<i>Speaking:</i> Maybe the learner should be informed and allowed to choose?

From the practical perspective of pronunciation training for second language learners, sticking to a single variety makes good sense. It would also be less work for the writers of teaching materials, and textbooks will be cheaper if they are shorter. At the same time, the capital-centered varieties are typically associated with high prestige. Would we not be doing second-language learners a favor if we teach them a prestigious variety? After all, most second-language learners do not have any clear affiliation with any of the target language regions.

However, it would probably be politically controversial to explicitly promote one variety over all the others. In addition, from the practical perspective of training listening comprehension, both to second-language learners and to young native speakers of another Scandinavian language, it is not smart the stick to one single variety. Non-native listeners will encounter variation in the speech of native speakers.

For our teaching material, we decided to represent at least two varieties from each language. Since there is an (possibly unconscious) idea of standard varieties, we wished for this variety to be represented. The choice of the other variety was less important because with this variety we above all wanted to give teachers an idea of how they can handle other varieties than the one represented in their textbook. If they wish to, they could then create their own complementary teaching materials for another variety.

Finally, we got recordings from two speakers of Southeast Norwegian (Oslo), and two speakers of Vennesla Norwegian (Southern Norway); four speakers of Copenhagen Danish, two speakers of Jutland Danish and one speaker of Funen Danish; and two speakers of Central Swedish (close to “standard pronunciation”) and two speakers of Scania Swedish (Southern Sweden). Occasionally, we also recorded some utterances for illustration ourselves, which added even more regional variety to the Norwegian speech sample.

5. Important vs. Less Problematic Features

Not all pronunciation features cause equally grave problems for interaction with native speakers. From the perspective of teaching, it would be desirable to determine what features are worth giving the greatest effort. For example, using the wrong tone accent for a word in Swedish and Norwegian will sound

“wrong” to native speakers, but it will probably not hinder comprehension. This makes sense, since there is already both variation in the pronunciation of tone accents between regional varieties as well as variation in the distribution rules for accent 1 and 2 to different words. This means that Swedish and Norwegian native speakers are already used to hearing some variations with respect to this feature. In contrast, stressing the wrong syllable in a word may cause real problems for comprehensibility [28].

Which features are thus relevant to teach, is essentially an empirical question. To our knowledge, there has not yet been much research in this area. Experienced language teachers probably have a pretty good idea of what features are the most important to get right for comprehensibility and for a good native-like pronunciation, even if this knowledge may be based on intuition and thus difficult to verbalize. It is desirable to research further what features are the most important to focus on in teaching, both for active pronunciation for language learners and for comprehensibility among native speakers of the Scandinavian languages.

6. The Prosody Visualization Tool

Visualization of intonation has proven to be helpful to language learners (see for example [29] for an overview over earlier research and a presentation of a tool developed in the 80s, and [30] for a test of our own Prosody Visualization Tool). In one of our projects, focus was on the development of an Prosody Visualization Tool that would be freely accessible and needed no or little training to use for teachers, language learner and everybody else interested in language and speech melody.

The tool can be used from different kinds of devices, but we have found it the easiest to use a computer with a headset. Language teachers and learners can speak into the microphone and see the visualization immediately and then save the output as a picture plus sound file. This online tool can assist teachers in creating teaching materials themselves, and students receive online feedback on their productions. Teachers who speak the target language well can also use the tool as a help for developing their own visualizations for teaching.

The visualization output is based on studies from a previous project. In that project, we were interested in finding an intonation visualization technique that was as intuitive as possible for second-language teaching. A stylized intonation contour, showing only the most important tonal turning points, along with stress markings yielded the best pronunciation [31]. The stylized intonation contour and the stress markings could be divided into two steps, or presented together. The Visualization Tool presents them together. In our teaching material, we used the Visualization Tool (sometimes also PRAAT [22]) to make handdrawn visualizations of intonation:

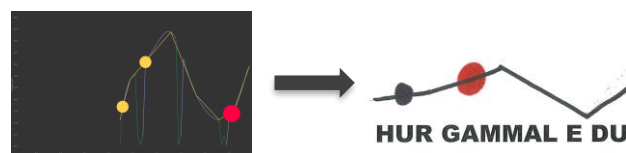


Figure 5: *To the left: output from the Visualization Tool of the Scania Swedish utterance ‘How old are you?’. To the right: a handdrawn visualization for teaching of the same utterance, based on the output from the Prosody Visualization Tool.*

Judging from the output of the tool, the most difficult part of the phonetic analysis is the identification of stressed syllables. This can be seen in Figure 5 above: the tool allocates one stress on *hur* ('how'), one on *gam(mal)* ('old') and then a bigger stress on *du* ('you') (actually, a nuclear accent in the terminology of Cruttenden [14]). However, if you listen to the utterance, you can clearly hear that there the stress on *gam(mal)* is the biggest accent of the sentence, whereas the last word *du* is completely unstressed. So, when using the tool, users need to use their own auditive judgement, but are supported by a visualization of the tonal movements and suggestions for the stress placement in the utterance. In opposition to PRAAT [22], for instance, the tool does not require a download, and it simplifies the presentation; compare the visualization produced by our tool with the PRAAT analysis of the same utterance as in Figure 5. The visualization in PRAAT in Figure 6 shows the analysis with all the default settings activated (left) and with only "Show pitch" activated (right); finding the right settings to display the pitch contour requires some rather advanced knowledge:

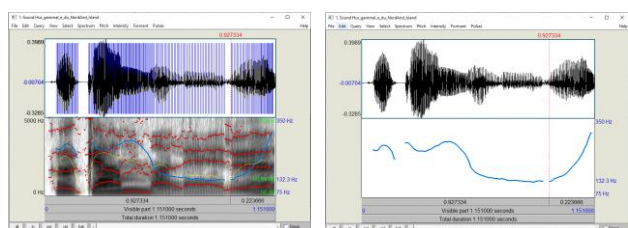


Figure 6: The PRAAT [22] visualizations of the Scania Swedish utterance *Hur gammal är du?* ('How old are you?'), also demonstrated in Figure 5.

Future research may show that there are adjustments that could be made to improve the Visualization Tool or some other tool similar to it. For example, when demonstrating our tool at conferences, we got the feedback from phoneticians working on Scandinavian languages that it would be useful to illustrate syllable duration because duration is one of the most important features of stressed and accented syllables in all three languages. In our tool, it is not possible to visually see the extension of the syllables because the tool does not parse the speech signal. It analyses the signal based on purely phonetic parameters. Adding a speech recognition component would be a project of completely different dimensions – especially since we want the tool to be usable for all varieties there are in the three languages.

Improvements that would make the tool really usable for language teachers would be to have a small video clip as output from the tool, so that sound and visualization are coupled. The possibility to upload own sound files into the tool is a prerequisite for using the tool for intonation training at home because the students cannot yet pronounce the target utterances themselves but need to upload a sound file with the correct pronunciation. For self studies it would also be necessary that the learner utterance was shown in visual comparison to the target utterance on the same screen [12].

Furthermore, it is possible that more fine-grained details of the tonal movements in accented syllables could be desirable for demonstrating lexically diverse tone accents. On the other hand, as we found out in our study [31], asking language learners to control too much tonal information at once may be confusing

instead of being helpful. In addition, in our material, tone accent movements are not always clear. In fast everyday speech, the accent melodies of the "smaller"/"non-focal" kind (see section 3) can be compressed or slurred until unrecognizability. The native speaker "hears" the accents anyway because the native speaker intuitively knows that there must be a tone accent on a certain word; but what is not there phonetically, the tool cannot show. Furthermore, tone accents are probably better explained and trained in another way, with focus only on the accent melody so that the learner can concentrate on the demanding timing of the accent movements. There are already some interesting methods, for example the small computer game "The Language Melody Game" (<https://projekt.ht.lu.se/lmg>) for learning to recognize (Central) Swedish tone accents [32].

7. The Teaching Material

In section 3, we presented the tonal parameters we decided to address in our teaching material. In section 4, we explained the considerations behind choosing varieties: the unofficial standard variety plus at least one other variety from each language. In section 5, we problematized what features should be spent most effort on. In section 6, we presented the visualization technique we use and which is also used as output in the Prosody Visualization Tool. The tool can be found here: <https://nordplus.sonoware.de/>.

In this section, we present some other pedagogical considerations that we have been discussing and that have influenced the design of our teaching material.

It was clear from the beginning that due to time limitations and copyright issues, we would not be able to go through films or video clips, for example on YouTube, to adapt for teaching purposes, even though videos with topics of social interest is probably the best way to capture the interest of teenagers. Instead, we decided to record our own audio material, in accordance with the GDPR.

Since we wanted the teaching material to be suitable for use in Swedish, Danish and Norwegian upper secondary schools, or even in the last two years of secondary schools, we have a target group of teenagers. Similarly, also university students, both native speakers studying their neighboring languages and exchange students coming in, are part of our target group. We would thus have preferred that the voices in our speech sample should be young, too, preferably with speakers between 15 and 25 years, to sound more interesting to the target groups. This proved difficult, though, as young teenagers would not lend their voices to our teaching material. To teenagers it matters immensely what their peers think of them, and they were probably afraid that the teaching material would not be considered cool by other teenagers. Therefore, most of our speakers are young adults, mainly university students, who are more self-confident. Only the two Scania Swedish speakers were under eighteen. They were part of the social network of one of the team members. But we still believe that our idea is right: if you create teaching material for young people, it should contain a good proportion of young speakers to have a chance of getting the other teenagers' interest.

We decided to use questions as examples. Questions are basic to communication. When you meet new people, you naturally start out by asking each other questions in order to get to know

each other. Or, if you are on vacation in a neighboring country, you may at some point have to ask some of the locals a question for finding a certain place, a good restaurant, how much something costs etc. Questions are also known to vary in intonation depending on the question type and communicative function of the question. We knew from the literature that Danish questions were supposed to work quite differently from Swedish and Norwegian ones with respect to intonation, using the declination of the utterance as a cue for its question type. Swedish questions will instead end with a high or low final boundary tone, more similar to English [14], see section 3. Even if the accent melody would end with a high or a low tone anyway, the tonal movement will rise or fall to an even higher/lower tone than usual at the end of a Swedish utterance. In Norwegian, final boundary tones are possible optional at least in some dialects, see Table 1, but the use of final boundary tones in different Norwegian dialects has not been systematically researched yet.

Another difference is the distribution of accent melody tones over the accent group. Swedish tone accent melodies are usually concentrated to the primary-stressed syllable of an accented word, maximally involving the first post-accent syllable as well to fulfil the accent melody. Swedish tone accents also seem to be limited to the lexical word in which they appear rather than being distributed over the whole accent group, which may contain other lexical words as well, see for example *RObotar ska* ('robots should') in Figure 3 above. This becomes especially clear when considering the last high tone in a Central Swedish higher-level accent 2 melody (**HL** + H [21]) as in *ÄLdrevården* ('care for senior citizens') in Figure 3 above. In long lexical words, typically compounds, the last H tone can be dislocated from the initial falling tone movement **HL**; but it still occurs in the same lexical word, more precisely on the last secondary-stressed syllable of that word [21, p. 124]. In contrast, the Norwegian accent melodies spread out over the whole accent group, often allocating the last accent melody tone on the last syllable of the accent group, which can be any lexical word and may be unstressed. So we expected questions to exhibit some interesting intonational differences across the languages.

We developed an interview script for our participants. We gave them the freedom to exchange words and reformulate if they thought it sounded more natural, but we presented them with a script, and most of them followed it rather closely. A script has the advantage of being controllable: We could ensure that several different question types were represented, and each question type more than once. Predetermined formulations kept the questions as parallel as possible with respect to syntax and word choice, which makes it easier for young native speakers to understand the questions in the neighboring languages so that they can concentrate on listening to the sound of them. Concentrating mainly on the scripted questions in the teaching material also makes participation less face-threatening for the participants: They do not give any personal information away. Furthermore, they are not responsible for the content of the questions and did not formulate the questions themselves.

We asked the participants to interview each other. This way, they spoke more naturally than if they had read questions aloud without having a conversation partner, and they were adapting the prosody to a real conversation, for actually getting an answer. The conversation partner answered the questions, and at times we can hear that the participants were really engaged in their

interview. This makes the questions more interesting to listen to and more authentic than if they were only read aloud from a list, without a conversation partner. It also makes the speech closer to everyday speech. From a scientific perspective, this is beneficial: This is how people in real encounters speak. From a pedagogical perspective, it can be problematic since it sometimes comprises fast, slurred speech and emotional emphasis, which does not always display the prototypical patterns desirable for teaching material.

Eventually, it became clear to us that we needed to start the teaching material with an introduction to what intonation is, before going into the tonal workings and differences. That is, we needed exercises in awareness raising (see section 2). In addition to the questions elicited, we also used some useful phrases as examples. For instance, many people enjoy being able to say "thank you" in several different languages. So we added the common phrases for greetings like "hi", "goodbye" and exclamations like "that's great!" or "are you out of your mind?!" These phrases were read off the list by our participants, without being integrated into interaction. Some of the phrases were selected for the material, presented with visualizations and a description of the most important pronunciation characteristics. The awareness-raising section is the only part of our teaching material in which we present utterances for active training. The rest of the material is aiming for familiarization by listening and pointing out intonational characteristics. The teaching material can be downloaded from the project's homepage: https://www.sdu.dk/en/om_sdu/institutter_centre/idk/forskning/projekter/human-robot+interaction/projects/nordplus-tool.

8. Summary and Outlook

In this paper, we have presented what we see as a road map to the development of comprehensive teaching materials for Danish, Norwegian and Swedish prosody. One of our aims is to make young native speakers familiarized with the sound of their neighboring languages. The other aim is to give language teachers a useful material for second language classes, as well as enabling them to make their own teaching materials if they rather want that.

Having team members from all three countries enabled us to start an intensive discussion of and comparison between the national prosodic traditions along with their terminology and definitions in all three languages. In the course of the project, we identified six characteristic tonal parameters and presented a first suggestion of terms and definitions for cross-linguistic comparison. However, much research remains to be done in this area; and even more research is needed in the area of what features are most worth the effort in pronunciation teaching.

It is not likely that Scandinavian teenagers will suddenly start spending time with the neighboring languages out of their own personal interest when they have not done so before, and so we aim at improving the teaching materials that school teachers and upper-secondary school teachers can use in school without having to put great personal effort into the teaching and without themselves having a good personal knowledge of or (initial) interest in the neighboring languages.

The best teaching material is one that evokes interest even if there was little or none to begin with. This suggests that the

material should deal with interesting topics and be activating and interesting to use. TV-programs and films should be made available to teachers, along with good speech corpora for both teachers and researchers. Freely available teaching platforms with ready-to-use teaching materials, including speech samples with ready-descriptions and explanations of different varieties should be created.

Even though our teaching material is just a small exploration of what could be done, it is freely accessible and can be used by teachers and learners to familiarize themselves or their students with the neighboring languages and prevent reactions like the impression that these languages sound “strange” or are “impossible to understand anyway”. With the visualization tool, we provide teachers with the opportunity to develop teaching material for active pronunciation training themselves, including material on non-standard varieties, by showing them how to use the tool and by using the visualization technique ourselves in our teaching material.

9. Acknowledgements

We gratefully thank the Nordplus Foundation for enabling the work presented in this paper by funding the two projects “Development of a Tool for the Visualization of Intonation” and “Developing Teaching Material for Speech Prosody to Increase the Mutual Comprehension between Danish, Swedish and Norwegian” from September 2020 to January 2021.

10. References

- Delsing, L.-O. & Lundin Åkesson, K. (2005). *Håller språket ihop Norden? En forskningsrapport om ungdomars förståelse av danska, svenska och norska*. TemaNord 2005:573. Copenhagen: Nordic Council. DOI: 10.6027/TN2005-573
- Gooskens, C. (2007). Contact, attitude and phonetic distance as predictors of inter-Scandinavian. http://www.let.rug.nl/gooskens/pdf/publ_nearcollateral_2007.pdf (Accessed on October 2, 2020).
- Bacquin, M. & Christensen, R. Z. (2011). När sommarfugle blir till sommarflugor – eller när grannspråk blir främmandespråk. *Sprog i Norden*, 2011, 157-166. <https://tidsskrift.dk/sin/article/view/17336/15061>
- Bohn, O. -S. & Askjær Jørgensen, T. (2017). A comparison of Danish listeners’ processing cost in judging the truth value of Norwegian, Swedish, and English sentences. *INTERSPEECH 2017*, 1741-1744. DOI: 10.21437/Interspeech.2017-9
- Språkrådet. (2019). Unge snakkar engelsk med danskar. <https://www.sprakradet.no/Vi-og-vart/hva-skjer/Aktuelt/2019/unge-snakkar-engelsk-med-danskar/> (Accessed on August 24, 2021)
- Nevile, M., & Wagner, J. (2011). Language choice and participation: Two practices for switching languages in institutional interaction. In: Pallotti, I. G. & Wagner, J. (eds), *L2 Learning as Social Practice: Conversation-Analytic Perspectives* (Volume 2, pp. 211-236). National Foreign Language Resource Center, University of Hawaii at Manoa.
- Einarsdóttir, S. B. (2017). Hvorfor forstår danskere ikke norsk? *Videnskab.dk*. <https://videnskab.dk/kultursamfund/hvorfor-forstaar-danskere-ikke-norsk> (Accessed on August 24, 2021)
- Nørgaard, F. (2019). Nabosprog som en integreret del af undervisningen. *Folkeskolen.dk*: https://www.folkeskolen.dk/838738/bachelor-nabosprog-som-en-integreret-del-af-undervisningen?utm_source=newsletter (Accessed on September 30, 2019)
- Henriksen, T. (2011). En ny tilgang til nabosprogsundervisningen. *Sprog i Norden*, 2011, 117-130. <https://tidsskrift.dk/sin/article/view/17334/15059>
- Grønnum, N. (2007). *Rodgrød med fløde: En lille bog om dansk fonetik*. Copenhagen: Akademisk Forlag.
- Grønnum, N. (2005). *Fonetik og fonologi: Almen og dansk*. Copenhagen: Akademisk forlag.
- Fischer, K., Niebuhr, O., Alm, M., Abelin, Å., Albertsen, E. & Asadi, A. (2021). Towards a Prosodic Visualization Tool for Language Learners. *1st International Conference on Tone and Intonation (TAI)*, 269-272. DOI: 10.21437/TAI.2021-55
- Alm, M., Abrahamsen, J. E., Abelin, Å., Albertsen, E., & Koreman, J. (2021). Parameters of tonal variation in and between three Scandinavian languages. *Proceedings of the 1st International Conference on Tone and Intonation (TAI), Sønderborg, Denmark*, 278-282. DOI: 10.21437/TAI.2021-57
- Cruttenden, A. (1986). *Intonation*. Cambridge: Cambridge University Press.
- Fretheim, T. (1990). The Form and Function of Foot-External Syllables in Norwegian Intonation. In: Wiik, K. & Raimo, I. (eds), *Nordic Prosody V. Papers from a Symposium* (pp. 87–110). Turku: Painosalama.
- Nilsen, R. A. (1992). *Intonasjon i interaksjon – sentrale spørsmål i norsk intonologi*. Dr.art. dissertation. Lingvistisk institutt, Universitetet i Trondheim.
- Fretheim, T. (1993). The Norwegian “Boundary Tone Agreement” Condition. In: Canakis, C. P., Chan, G. P. & Marshall Denton, J. (eds), *Papers from the 28th Regional Meeting of the Chicago Linguistic Society 1992. Volume I: The Main Session* (pp. 159–170). Chicago: Chicago Linguistic Society.
- Kristoffersen, G. (2000). *The Phonology of Norwegian*. Oxford & New York: Oxford University Press.
- The Trondheim Model of Intonation, developed by Thorstein Fretheim and Randi A. Nilsen, see [18: Ch. 10].
- Molnár, V. (1991). *Das TOPIK im Deutschen und im Ungarischen*. Stockholm: Almqvist & Wiksell. Ch. 4.
- Myrberg, S. & Riad, T. (2015). The prosodic hierarchy of Swedish. *Nordic Journal of Linguistics*, suppl. Prosody in the Nordic Languages, 38(2), 4-147. DOI: 10.1017/S033258653000177
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International* 5(9/10), pp. 341-345.
- Niebuhr, O. (2021). The Kiel Intonation Model – KIM. In: Barnes, J. & Shattuck-Hufnagel, S. (eds.), *Prosodic Theory and Practice* (pp. 287-318). MIT Press.
- Hvad er forskellen på sprog og dialekt? Copenhagen University, dialekt.dk. https://dialekt.ku.dk/dialekter/sprog_standard_og_dialekt (Accessed January 19, 2022)

25. Lund, J. (1984). Den danske skole og dialekterne: En studie i sejlads uden ror. *Sprog i Norden* 1984, pp. 19-35. <https://tidsskrift.dk/sin/article/view/17648/15383> (Accessed February 25, 2021)
26. Vad är rikssvenska, när uppstod den och var kommer den ifrån? *Frågelådan*, Swedish Institute of Language and Folklore. https://frageladan.isof.se/visasvar.py?sok=rikssvenska&sv ar=44972&log_id=852856 (Accessed October 13, 2021)
27. Høyland, B. (2021). Lærarspråk i norsk som andrespråk når nynorsk er opplæringspråket. *Språkprat* 3, December 2021. <https://sprakprat.no/2021/12/03/laerarsprak-i-norsk-som-andresprak-nar-nynorsk-er-opplaeringspraket/>
28. Abelin, Å. & Thorén, B. (2016). Identification of stress, quantity and tonal word accent in Swedish. Book of abstract of the Conference *New Sounds* 2016, Aarhus University. https://conferences.au.dk/fileadmin/user_upload/New_Sounds_2016_Book_of_Abstacts_internet_version.pdf (Accessed on May 29, 2022)
29. Weltens, B. & de Bot, K. (1984). The visualization of pitch contours: Some aspects of its effectiveness in teaching foreign intonation. *Speech Communication* 3(2), 157-163.
30. Fischer, K. & Niebuhr, O. (2022). The Effects of the Online Visualization of Acoustic-Prosodic Features of Speech on Speakers' Productions. *Proceedings of the 33rd Conference on Electronic Speech Signal Processing*. Sønderborg, DK.
31. Niebuhr, O., Alm, M., Schümchen, N. & Fischer, K. (2017). Comparing visualization techniques for learning second language prosody: First results. *International Journal of Learner Corpus Research* 3:2, 250–277. DOI: 10.1075/ijlcr.3.2.07nie
32. Schremm, A., Hed, A., Horne, M., & Roll, M. (2017). Training predictive processing with a digital game: Prototype promotes acquisition of anticipatory use of tone-suffix associations. *Computers & Education*, 113, 206-221. DOI: 10.1016/j.compedu.2017.07.006