Towards a Prosodic Visualization Tool for Language Learners

Kerstin Fischer,[1] Oliver Niebuhr,[1] Maria Alm,[1] Åsa Abelin,[2] Egil Albertsen[3] & Ali Asadi [1]
[1]University of Southern Denmark
[2]Gothenburg University
[3]Norwegian University of Science and Technology
[kerstin, olni, mhalm]@sdu.dk, asa.abelin@ling.gu.se, egil.albertsen@ntnu.no

Abstract

In this paper, we present an online tool for the visualization of prosodic information developed in a Nordplus project to enable speakers of the Scandinavian languages to understand their neighboring languages better. The free online tool visualizes intonation contours in different degrees of stylization and presents stress marks for syllables that are emphasized; furthermore, it provides different download and playback options. We describe the steps we have taken to ensure that the tool is useful for language teachers; specifically, we report on a qualitative user study in which 14 teachers have tested the tool and its functionalities. We can conclude that the visualization tool is likely to facilitate the creation of new teaching material for teachers and possibly also for language learners.

Index Terms: speech prosody, visualization, language learning

1. Introduction

In this paper, we describe our steps towards the development of a web-based tool that provides online visualizations of the prosodic realizations of utterances. The tool development was carried out in the framework of a Nordplus project, whose aim is to support the mutual intelligibility between the Scandinavian languages. Specifically, we developed a free online tool to help young speakers understand the prosodic differences between the three languages by visualizing the relevant prosodic features and distinctions in an understandable way. This should enable speakers of one Scandinavian language to comprehend the other Scandinavian languages better. We furthermore expect the tool to allow teachers to create teaching materials themselves, as well as students to practice themselves. We thus address teachers without special training in prosody and create exemplary teaching material for teachers to download. In the present paper, in order to verify that the tool is helpful and understandable also to users without training in prosody, we carried out think-aloud usability testing with language teachers with different backgrounds.

2. Tool Design Requirements

An initial version of the visualization tool was created against the background of years of experience with software-supported public-speaker training [9] and our previous studies of different prosodic notation systems [8]. In the latter, we found iconic, i.e. visual, notation systems to be easier to interpret by language learners than symbolic ones. Furthermore, we found that stylized contours are easier for users to interpret than concrete and detailed ones. Specifically, we found that the stylized "hat pattern" leads to the fewest errors and most positive evaluation, and that all visualizations are preferred over the symbolic notation systems.

In a more recent paper, we explored under what circumstances information on prosodic stress and intonation can be presented together, because our initial investigation suggested that notation systems that visualize the intonation contour only are easier and more intuitive and lead to fewer production errors than notation systems that also include information on stress.

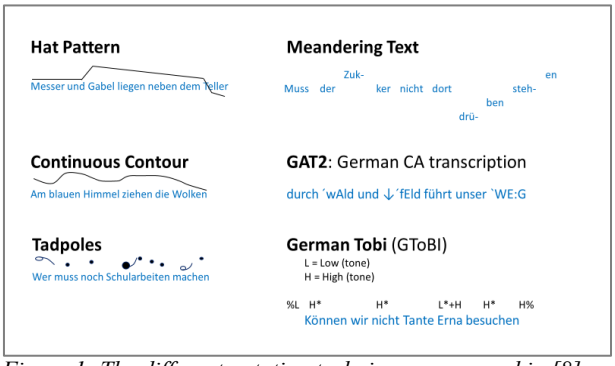

Figure 1: The different notation techniques compared in [8]

In our follow-up study (Fischer et al. [3]), we found that adding information on stress is both perceived as beneficial by language learners and leads to fewer mistakes, but that it should be combined with stylized contours and possibly presented in an additional step.

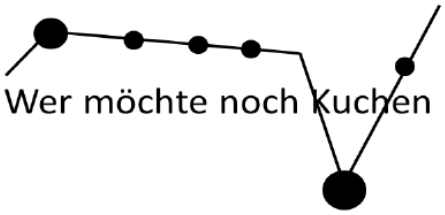

Figure 2: The best performing visualization comprises stylized contours with stress marks [3]

The study results together suggest that the tool to be developed should support stylization of intonation contours, present stress patterns, and allow for separate presentations of intonation contours and stress. In addition, the tool should of course comply with general usability criteria and be intuitive and easy to use by untrained users.

3. The Visualization Tool

The visualization tool for language teachers and learners developed based on these findings produces visualizations of

10.21437/TAI.2021-55

intonation contours with different degrees of stylization. The tool allows teachers to receive online visualizations of their utterances, so that they can create their own teaching material. It provides the opportunity to visualize the raw data, the unsmoothed contour, stress markers and the stylized contour, where the degree of stylization can be adjusted. Sections of creaky voice are separately highlighted as well, so that learners can see and practice the variation in voice quality that coincides in language-specific ways with turn-taking and backchanneling signals or with certain tone and intonation patterns.

Furthermore, the tool allows the user to replay the audio at different speeds and to save both audio and images. The intended use of the system consists of two scenarios: A language teacher uses the tool to produce target utterances, to save the audio files and images of the visualizations of the utterances and to present them to language learners according to his or her wishes and in the local dialect to be taught to the students. In a second, related scenario, the learners can use the tool to produce their own visualizations and to compare them to those of the teacher.

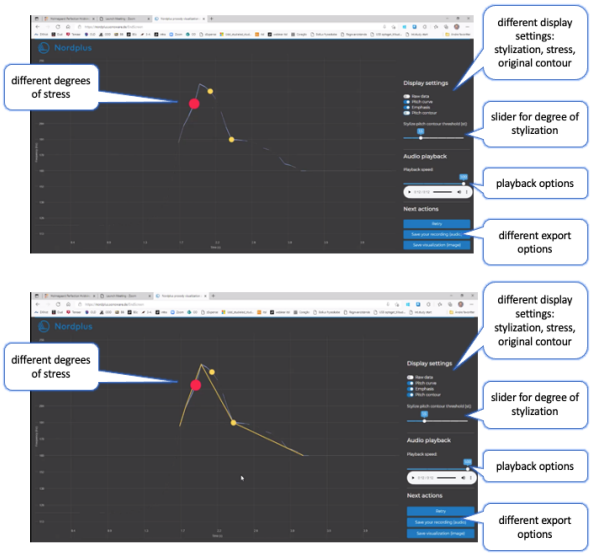

Figure 1: *The visualization tool with different degrees of stylization displayed*

The tool can be found under:

"https://nordplus.sonoware.de/"

4. User Study

The aim of the user study was to identify teachers' specific needs with respect to prosody teaching, to understand their current practices and circumstances of teaching, and to get feedback on the visualization tool.

4.1. Method

In a user study, we carried out interviews with language teachers of the Scandinavian languages, as well as with some additional teachers who have specific experience with pronunciation and prosody teaching. 11 interviews were carried out online via zoom, 3 were carried out in written form.

4.2. Procedure

First, we asked participants to fill out a consent form, after which we started the recording. Our first questions concerned teachers' current teaching practices concerning intonation teaching. Specifically, we asked them about their current practices concerning the teaching of intonation, what exactly they teach and what they would wish for to support their teaching. Then we invited them to open and try out the tool and to verbalize what they were thinking while encountering the tool and its functioning.

During the usability tests, we had participants share their screens and open the browser window with the visualization tool. However, due to the online interview situation, technical problems occurred for several participants concerning microphone use; they had the choice of either remaining in the online zoom meeting or try the tool. In these cases, the interview was interrupted, so that participants could try the tool themselves, and then continued afterwards. In this way, participants comments and evaluations were recorded, but not their ongoing thought processes and potential usability problems that characterize think-aloud protocols (cf. [7]). Nevertheless, as Krug [6] states, 3-5 participants are usually sufficient to identify the most crucial usability issues in a think-aloud usability test, so that our user study is likely to have revealed the most important pitfalls of the tool.

The think-aloud usability tests were facilitated using a conversational approach (cf. [5]) in order to mitigate the perceived awkwardness of the situation for our participants, who were not used to the think-aloud method, and in order to encourage them to try different functionalities of the tool [2].

The different interview sessions were partly transcribed and analyzed by creating initial codes, which were then grouped thematically. In the next step, supporting quotes were identified and translated since interviews were carried out in five different languages, and the thematic analysis was iteratively revised.

4.3. Participants

We carried out 12 semi-structured interviews with subsequent. think-aloud usability testing ([2]; [6]). One person refused to be recorded, so we excluded her data from the subsequent analysis, leaving 11 valid interview participants. The participants were recruited based on personal connections by the different authors and comprise:

- 4 participants who teach Danish as a second language
- 1 participant who teaches Danish to Germans
- 1 participant who teaches Danish pronunciation to Scandinavian native speakers
- 2 participant who teaches Danish to immigrants and international students
- 3 participants who teach Swedish as a second language to immigrants or international students
- 4 participants who teach English as a foreign language in Iran

In addition, 3 participants who teach Norwegian as a second language participated via an e-mail interview, in which the users tried and assessed the tool independently.

4.4. Results

The results concern the teachers' current practices regarding the teaching of prosody, their evaluation of the tool in the think-aloud usability tests, and their suggestions for improvement.

4.4.1. Current Teaching Practices

Before we exposed the participants to the visualization tool, we asked our interview participants about their current practices in teaching intonation; of our 14 participants, only 11 teach intonation at all. One teacher stated: "I teach intonation but not as a separate sub-skill. I integrate it with other skills." Several teachers stated that there is generally not enough time for intonation. Six participants said it is important to teach stress and intonation in classes. For example, one said, "In Danish, it is important to teach basic intonation and stress structure of phrases. They are rule-based, and we should teach the language melody". Another teacher said, "stress and intonation go together, so you cannot separate them". One interviewee also said it is important to teach intonation "when you want to master the flow of speech". One participant also acknowledged that intonation may be somewhat different, but is not very problematic.

The methods that they use comprise:

- modeling, i.e. the oral presentation of the target realization
- repeating in chorus
- using gesture to indicate up and down movement of the voice
- using a musical instrument
- drawing arrows or contours themselves. For instance, one participant answered: "I write sentences on the board and draw arrows to show their intonation patterns."
- introducing stress rules: "stress and intonation go together"
- using textbooks, like 'rød grød med fløde', or teachers' own material

When asked which phenomena they focus on, they replied:

- sentence intonation, especially questions
- accent 1 and accent 2
- stress and intonation combined
- pronunciation, then intonation follows naturally

We also asked them what they would like to have to support their teaching. Their replies comprise:

- (engaging) authentic examples
- a musical instrument
- visualization of the contours
- exercises

Specifically, one teacher stated that "authentic examples could be good for most students" and another pointed out that "video files and movie clips that engage students in the lesson are helpful." Three teachers said that visual representations and sound files would facilitate the teaching of intonation. One teacher suggested that a "visualization program which is as fast as the one we are watching" would be helpful for students to have. Two teachers said they use text-based materials and exercises such as the Danish sentence "rød grød med fløde". One teacher said he uses his hands to show rising/falling intonation, and another one said that she'd "write sentences on the board and draw arrows to show their intonation patterns."

4.4.2. Evaluation of the Tool

10 out of 14 participants found the tool generally useful: "t's a good visual aid, also in online classes. I really like the tool and want to use it in my classes." One participant said that she finds the tool easier to use than PRAAT (a prosodic analysis tool, cf. [1]), and she especially mentions the stress dots, which PRAAT cannot do. Moreover, she finds the playback function with the possibility to slow speech rate down also useful. Furthermore, ten participants mentioned explicitly that they liked that they could record sentences. These participants also thought that the tool was easy to use.

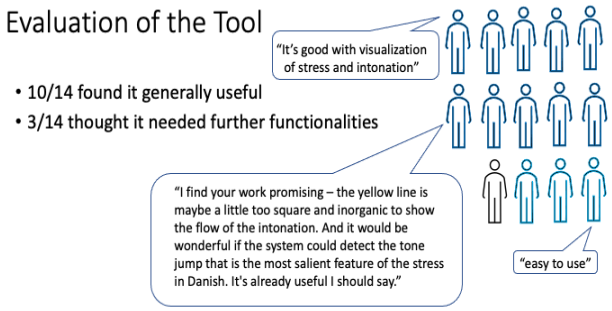

Figure 2: *General evaluation of the tool*

Some participants faced usability issues; for instance, one said that it "took a little time before I understood that I should speak." One participant experienced a lag in the tool as it was a bit slow. Another participant took a long time to find the stop button and didn't immediately get the playback function and the re-record functions. One teacher also mentioned that she was uncertain what the difference between pitch curve and pitch contour was.

Moreover, for three participants, the tool only recorded half (or less) of the sentences/words they uttered. It was not as sensitive as they expected, which made them uncomfortable. In addition, some minor usability issues showed up, like problems regarding the labels of the different display options and the length of the welcome blurb, which most perceived as too long. In sum, the usability issues identified comprise:

- the welcome blurb is too wordy
- it is unclear what the start-button does (record)
- some labels are unclear
- the tool is possibly not sensitive enough with some microphones

Several participants also commented that the stylized contours were too stylized, for instance: "it's too angular", without realizing that they can actually choose the degree of stylization in the tool.

Furthermore, some additional needs could be identified, like the need to zoom into the representation. One participant said he would like to mark parts of the resulting curve, zoom in on it, and listen over and over. Another participant said he would like to compare target utterance and learner utterance through two visual windows, preferably also having students upload their target files into the tool at home. Another participant would like to upload two teacher utterances to

compare at the same time. Thus, participants made some suggestions for improvement:

- many teachers thought it useful to be able to see two contours at the same time: the teacher's and the student's
- they asked for additional functionalities, e.g.
  - the possibility to download videos
  - the possibility to zoom in
  - the possibility to align the visualization with recognized speech
- an "oscillogram or spectrogram in order to see or show which syllables or words the tone curve is associated with"
- the ability to upload speech samples

Furthermore, it also became also apparent that some teachers did not know what they would do with a tool like this: "I think I'll need training on how to use the tool". Many participants did not come up with ideas on how to employ the tool themselves, and many would not have actively looked for such a tool.

5. Discussion

Almost all interviewees found the tool to be potentially very useful and embraced its intended uses. However, when analyzing the interview and usability testing results, we have to take into account that participants were aware that their interviewers were involved in the development of the visualization tool, which may have led to some courtesy bias.

Several usability issues could be identified, some of which could be fixed easily (like the welcome blurb). Others depend on computer hardware (like the sensitivity of the in-built microphones) or on the network connection (like delay). Regarding additional functionalities, the study revealed opportunities for future work, which would make the visualization tool even more useful.

Furthermore, it became clear that some teachers needed additional information, like ideas on how to use the tool in class. Thus, even though the aim of the tool is to allow teachers to create their own, localized and context-sensitive teaching material, some example uses and exemplary teaching material to accompany the introduction of the visualization tool can be useful.

6. Steps Taken and Future Work

To address the usability issues concerning the labels of the different display options and the participants' lack of imagination on how to use the tool in their classes, we created introductory teaching material and example exercises to give teachers a better idea of how the tool can be used. Specifically, we created a one-minute video that introduces teachers to the use of the tool. This video can be found here:

"https://www.sdu.dk/en/om_sdu/institutter_centre/idk/forskning/projekter/human-robot+interaction/projects/nordplus-tool/teaching-materials/tool/tutorial"

Moreover, we created extensive teaching material (79 pages, 82 examples, 157 sound files), which displays target contours, which can serve both teachers and learners as learning targets to which they can compare their own contours. This teaching material can be found here: "https://www.sdu.dk/en/om_sdu/institutter_centre/idk/forskning/projekter/human-robot+interaction/projects/nordplus-tool/teaching-materials".

The implementation of aligned windows, the integration of speech recognition software and further download and viewing options will be targeted in future work.

Since the visualization tool was developed specifically with the Scandinavian languages in mind, the prosodic features visualized focus on what is relevant in these languages; however, we found the tool also to be helpful for speakers of German [4], who were asked to produce utterances with deep falling final contours, using the visualization of creaky voice of the tool. The study shows that the visualization improved participants' productions significantly. Thus, the visualization tool was found to facilitate learner productions also in German. Nevertheless, it is open to what extent the tool facilitates learner productions in languages that rely on very different prosodic information.

7. Acknowledgements

This research was funded by Nordplus in the framework of 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*,' which we gratefully acknowledge; the authors would also like to thank the teachers who participated in the interviews and questionnaire study.

8. References

[1] P. Boersma. 2001. Praat: A system for doing phonetics by computer. *Glot International* 4, 341-345.

[2] M. Boren & J. Ramey. 2000. Thinking aloud: Reconciling theory and practice. *IEEE Transactions on Professional Communication,* 43 (3), 261-78.

[3] K. Fischer, Niebuhr, O., Alm, M. and Schümchen, N. submitted. Intuitive Visualizations of Intonation for Foreign Language Learners.

[4] K. Fischer & O. Niebuhr, 2022. The effects of the online visualization of acoustic-prosodic features of speech on speakers' productions. In O. Niebuhr, M. Svennson-Lundmark, & H. Weston (Eds.), *Studientexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung 2022* (Vol. 103, pp. 180-187). TUDPress. Studientexte zur Sprachkommunikation https://www.essv.de/paper.php?id=1157

[5] M. Hertzum. 2016. A Usability Test is not an Interview. *Interactions of the ACM*. Evaluation and Usability Forum.

[6] S. Krug. 2014. Don't Make me Think. Indianapolis: New Rider.

[7] L. Lentz & H. Pander Maat. 2007. Reading Aloud and the Delay of Feedback: Explanations for the Effectiveness of Reader Protocols. *Information Design Journal* 15, 3: 266-281.

[8] O. Niebuhr, 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).

[9] O. Niebuhr and Neitsch, J. 2020. Digital Rhetoric 2.0: How to Train Charismatic Speaking with Speech-Melody Visualization Software. In: Karpov A. & Potapova R. (eds), *Speech and Computer (SPECOM 2020). Lecture Notes in Computer Science, vol 12335* (pp. 357-368). Bern: Springer, Cham.