Dealing with L1 background and L2 dialects in Norwegian CAPT

Olaf Husby[1], *Åsta Øvregaard*[1], *Preben Wik*[2], *Øyvind Bech*[1], *Egil Albertsen*[1],
Sissel Nefzaoui[1], *Eli Skarpnes*[1], *Jacques Koreman*[1]

[1]Department of Language and Communication Studies, NTNU, Trondheim, Norway
[2] Department of Speech, Music and Hearing, KTH, Stockholm, Sweden

olaf.husby@ntnu.no, asta.ovregaard@ntnu.no, preben@speech.kth.se, oyvind.bech@gmail.com,
{egil.albertsen, sissel.nefzaoui, eli.skarpnes, jacques.koreman}@ntnu.no

Abstract

This article describes the CALST project, in which the primary aim is to develop Ville-N, a CAPT system for learners of Norwegian as a second language. Since there is no accepted pronunciation standard in Norwegian, the system uses four dialects (represented by one male and one female speaker each). Ville-N makes use of L1-L2*map*, a tool for multi-lingual contrastive analysis, to generate a list of expected pronunciation problems. These can be used to tailor pronunciation and listening exercises. The tool can also be used for other target languages. We propose L1-L2*map* as a collaborative tool for the CAPT community.

Index Terms: CAPT, Ville-N, Norwegian, dialects, multi-lingual contrastive analysis, L1-L2*map*

1. Introduction

CALST (Computer-Assisted Listening and Speaking Tutor) is a project headed by the Norwegian University of Science and Technology (NTNU) in collaboration with the University of Oslo (UiO), The Adult Education Centre (EVO) in Trondheim, and the Royal Institute of Technology (KTH), Stockholm (Sweden). The project's aim is to develop a computer-assisted pronunciation training (CAPT) system for Norwegian as a second language (NSL). Two features of the system are important:

1) The CAPT system allows the selection of four different dialects of Norwegian, with one male and one female speaker to represent each dialect.

2) The CAPT system makes use of L1-L2*map*, a tool for contrastive phonological analysis for all relevant L1-L2 pairs. The tool is used to better tailor exercises to the needs of the Norwegian learner, but can also be used by CAPT developers for other target languages.

The CAPT system for Norwegian is based on Ville, the virtual language teacher for Swedish [1]. The Norwegian system, which was named Ville-N, makes use of a database consisting of "1000 words and expressions" divided over 43 semantic or syntactic categories which the language learner can select, such as "food", "education", "animals" or "pronouns". The words and expressions were selected on the basis of their relevance to beginning learners of Norwegian and correspond to the A1 and A2 levels in [2]. For each word, a corresponding picture is stored. About 30% of the pictures are taken (with permission) from [3], and these are complemented by drawings made by a local artist in the same artistic style in order to get a consistent and coherent set of images. For each word, the article and derivations (e.g. "en bil, bilen, biler, bilene" for "a car, the car, cars, the cars") or inflexions ("å gi, gir, ga, gitt" for "to give, give, gave, given"), as well as an English translation and the grammatical class

(noun, verb, etc.) are stored and displayed on the screen during basic vocabulary training as useful information for the learner. The basic CAPT system consists of three different parts:

1) Perception training for "1000 words and expressions" which are pronounced by a virtual language teacher or talking head, while a corresponding 2x2, 3x3 or 4x4 picture grid is shown on the computer screen (Figure 1).

2) Pronunciation training where the language learner can compare his/her pronunciation of the words with that of the virtual language teacher. Although the system does not analyze the learner's pronunciation, the language learner is given the opportunity to develop greater awareness of pronunciation problems through self-monitoring.

3) "Listen and write" spelling exercises where the language learner listens to the 1000 words and expressions and types them in.

In parts 1 and 3 the virtual language teacher gives feedback on the correctness of the learner's answers. The basic Ville system is described in more detail in [1].

First pilots with Ville-N have been carried out in the Norwegian courses for foreign students and employees at NTNU. The system will also be used for the same target groups at the University of Oslo, as well as in several courses for teachers of Norwegian as a Second Language. In addition, the collaboration in the project with EVO widens the target group to L2-learner groups that include illiterate learners and learners from languages and social backgrounds different from what usually is found at the universities. This is of interest because it puts high demands on the user interface.

In the following section we describe the specific situation for language learners who want to learn Norwegian and the consequences this has for the design of the CAPT system. In Section 3 the tool L1-L2*map* is described. This is a multi-lingual tool which can produce a contrastive analysis for a large number of source languages (L1's) compared to Norwegian as the target language (L2), but it can also be used by CAPT developers for other target languages. We discuss the consequences of using a multi-lingual contrastive analysis, and describe how the tool is embedded in the CAPT system.

2. Norwegian CAPT

There is no accepted standard pronunciation for Norwegian. This poses a challenge to learners of Norwegian. The use of a CAPT system in NSL teaching supports language learners in dealing with the variations in pronunciation of Norwegian. In Sections 2.1 and 2.2, the (written and spoken) language situation in Norway and the ways in which it differs from that in many other countries is discussed. In Section 2.3 the teaching practice in Norwegian language classes is described, and in Section 2.4 some of the consequences for the design of the CAPT system are discussed.

2.1. The written standards of Norwegian

There are two written forms of Norwegian: *Bokmål*, which is used by approximately 85% of the population, and *Nynorsk*, used by about 15%. While *Bokmål* was developed from Danish and Norwegian spoken by the "educated" classes in major cities, the orthography for *Nynorsk* was based on Norwegian dialects and today has its point of gravity in the southwestern part of Norway. *Bokmål* is the dominating written form elsewhere, and was therefore selected for the orthography in the first version of Ville-N.

There is no *spoken standard* for either of the two variants of Norwegian, but as one would expect there are dialects/sociolects that are closer to one or the other written form.

2.2. Dialectal variation

Traditionally Norway is divided into four major dialect areas which can be subdivided further [4]. Northern and Mid-Norwegian dialects have a rather rich consonant phoneme inventory including retroflex and palatal sounds. The two dialect groups are divided by prosodic features (high vs. low tone on stressed syllables). Southeastern dialects (which are low-tone dialects) lack palatals. Western dialects lack both palatal and retroflex sounds, the latter due to the fact that these dialects use a dorsal /ʀ/ instead of an apical /r/ which triggers retroflexion.

Since there is no accepted standard pronunciation of Norwegian, speakers use their dialect in widely varying contexts, with only a weak tendency towards normalization (or use of a pronunciation variant closer to either of the two written standards). Generally, dialects have a high status and are used at all levels of Norwegian society, both in public meetings, in churches, in parliament, on television and on the radio. Norwegian language users (are considered to) understand all the varying forms of spoken Norwegian.

The lack of a spoken standard also leads to the consequence that Norwegians in general are poorly skilled when it comes to speaking forms of Norwegian that reflect the written forms. Norwegians are not given formal instruction in how to speak Norwegian. In school one may observe that some students, while reading Norwegian texts aloud, follow the orthography, but accommodate their reading to the phonological pattern characterizing the dialect that they normally use. The widespread use of dialectal speech puts the burden of dealing with dialectal variation in everyday communication more with the language learner than in the case of languages that do have a standard pronunciation.

2.3. Teaching Norwegian to foreigners

Most teaching material in Norwegian as a second language is written in *Bokmål*. This material is also commonly used in areas in which *Nynorsk* is the dominating written form. Teachers of Norwegian as a second language (NSL) usually accommodate their speech to *Bokmål* in order to minimize the difference between the spoken form and the written form in most textbooks. To some degree they adapt the phonological inventory of southeastern Norwegian, while they often maintain the prosodic features of their dialect. In this aspect, teachers of NSL represent a communicative practice that is rare among Norwegians. Students attending NSL courses all over Norway are therefore acquiring a form of Norwegian that is not commonly used outside the classroom.

This puts the NSL speaker in a special position as they speak a form of Norwegian that is not used locally, and as listeners they are not tuned to the locally spoken forms. In order to communicate effectively, NSL speakers should learn to speak one dialectal form of Norwegian (possibly the local variant of the area where he/she lives), but understand all dialects. One former student of NSL claimed that he needed two years to acquire Norwegian: one year to learn to read, write and speak, and another year to learn to listen to *trøndersk*, the dialect of in the middle part Norway (which is characterized by special features such as lowered front vowels, vowel harmony, retroflex and palatal sounds, apocope, and special properties related to morphology, syntax and lexicon). It is difficult for teachers (unless they are expert dialect imitators) to teach L2 learners different dialects, and Ville-N thus offers a unique way to deal with this problem.

2.4. Implementation in the Norwegian CAPT system

In order to accommodate to the linguistic situation in Norway, the pronunciation of each word by one male and one female speaker in each of the four Norwegian dialects in Section 2.2 is stored in the database of "1000 words and expressions".

2.4.1. Multiple speakers

Language learners can either select a single speaker (talking head) in Ville-N or choose to listen to several dialects in random alternation. Exposure to more than one voice is generally useful for language learning [5]. For familiarization with different dialects, it is important to listen to different speakers in Ville-N, both male and female, pronouncing the words in different dialects. To enhance learning, we have included a hyperarticulated pronunciation of each word (which can be accessed via a "Repeat utterance" button), while a normal, casual pronunciation is used otherwise. For practicing pronunciation, on the other hand, it is useful to select a single role model, preferably a speaker of the same sex as the learner.

2.4.2. Homonyms

Some words that are different in one Norwegian dialect are homonyms in another. For example 'yours' and 'theirs' are both 'deres' in southeastern dialects, but 'dokkers' and 'dis' in western Norway. This means that while there is a one-to-one correspondence between word and picture in western Norwegian, there are two pictures corresponding to 'deres' in southeastern dialects. Here, the learner must get positive feedback from the virtual teacher when he/she clicks on either of the two pictures corresponding to 'yours' and 'theirs' (cf. Figure 1).
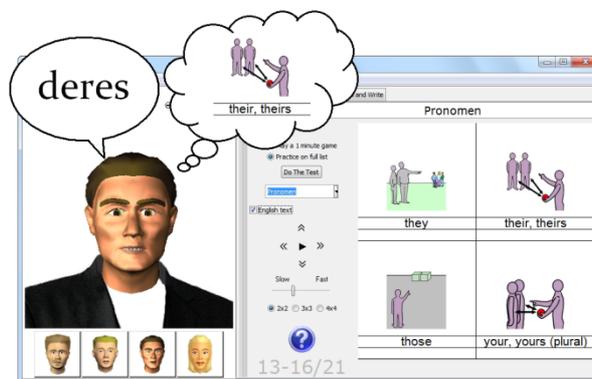

Figure 1: *An example of a homonym in one dialect, but not in another. 'Theirs' and 'yours' (pl) are the same in southeastern Norwegian, but not in other dialects.*

3. L1-L2*map*

The basic Swedish CAPT system Ville has been extended with exercises which allow the language learner to train certain phonemic oppositions by listening to Swedish minimal pairs demonstrating for example the vowel length or plosive voicing opposition. We are now in the process of implementing similar exercises which have already been developed for Norwegian, and a simple phonetic explanation of each opposition will be provided to the language learner together with each exercise.

But since the required listening and pronunciation training depends on the language learner's native language, it can be useful to carry out a contrastive analysis for the source (native) versus target language, in our case Norwegian. On the basis of the result of the contrastive analysis, a speaker of standard Italian will be directed to an exercise on short versus long vowels in Norwegian, since he/she only has short vowels in his/her native language and may therefore experience difficulty with long vowels. A German speaker will not take this exercise, because German also has a vowel length opposition and the learner is therefore familiar with the phenomenon of vowel length.

We use L1-L2*map* [6] to carry out an automatic contrastive analysis where the source language can be chosen from a large number of languages, with (one of the dialects of) Norwegian as the target language. L1-L2*map* is based on the UPSID database, which contains 451 languages [7]. This number has been increased to more than 500 languages in L1-L2*map*. As shown in Figure 2, the result of the contrastive analysis is shown in four tabs (comparable to the tabs in your browser). The choice of tabs is based on the IPA representation [8]: "Consonants", "Consonants (other)", "Vowels" and "Diphthongs". The "Consonants" tab presents pulmonic consonants, while non-pulmonic consonants and affricates are presented in "Consonants (other)". The latter is only shown if relevant for the L1-L2 comparison. For the sake of simplicity rows in the consonant tables (representing manner of articulation) are only visualized if used in at least one of the languages. A fifth, "Language information" tab presents some general information, following the information given in UPSID. A lay-out which is very similar to that used in [8] was used in order to provide language experts using the system with a simple and recognizable lay-out.

The colours in L1-L2map have easily interpretable functions: red indicates sounds that only occur in the target language (cf. red language box at the top of the window); blue indicates sounds that are only used in the learner's native language (cf. blue language box); and green indicates sounds that L1 and L2 have in common. It is the sounds which are indicated on a red background which the CAPT system needs to focus on in listening and pronunciation exercises for the language learner.
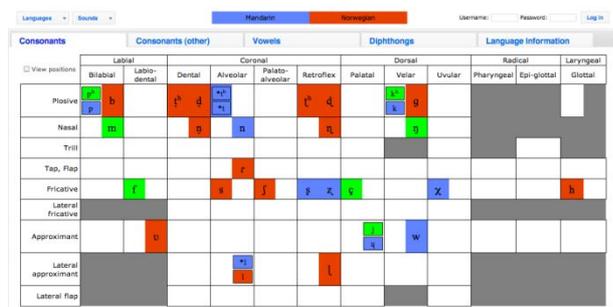

Figure 2: *Lay-out of the result of a contrastive analysis comparing Mandarin Chinese with (southeastern) Norwegian*

In the following sections, we first describe how a contrastive analysis can generate a list of possible pronunciation problems for the language learner (Section 3.1). In Section 3.2, we discuss the effect of the level of granularity of the analysis, and how this is determined by our choice of UPSID as the basis for a contrastive analysis. In Section 3.3 we explain how L1-L2*map*, the tool for contrastive analysis [6], is embedded in the CAPT system. In Section 3.4 we invite language experts to contribute to and use the tool.

3.1. Contrastive analysis

The contrastive analysis hypothesis (CAH) attempts to predict where language learners may encounter difficulties when acquiring a target language. Lado states that a comparison of two languages can be used to "predict areas that will be either easy or difficult for learners" [9] and that problems in learning a foreign language can be explained by transfer phenomena induced by the learner's native language. Today we know that L1 transfer is far from the only factor affecting pronunciation of an L2, but it is still a starting point when developing teaching material which aims to support students in acquiring a more intelligible pronunciation [10]. It is important to bear in mind that the process is two-fold as phenomena related to both perception and production are at play. One issue is related to how L2 learners can overcome the L1 phonological filter in their perception of a foreign language, the other is how to modify L1-related motor patterns and establish new patterns that are adequate for L2 pronunciation.

When creating exercises for second language learning, a contrastive analysis may bring to light different types of problems. The first three are taken from [11]:

1) *Neither phoneme of a foreign language minimal pair exists in the native language.*

   The Norwegian /yː/-/ʉː/ contrast is difficult to hear and produce for most learners of Norwegian. In general the sounds are replaced by /iː/ and /uː/, respectively. The best strategy for learning the phonemes is to first train each of them in another minimal pair, in this case, /iː/-/yː/ and /ʉː/-/uː/, before training the /yː/-/ʉː/ contrast.

2) *One of the phonemes of a minimal pair exists in the learner's native language.*

   The Norwegian /ʃ/-/ç/ contrast is difficult for most learners of Norwegian. In general the palatal fricative /ç/ is out of reach for beginners and often replaced by /ʃ/.

3) *Both speech sounds exist in the native language, but only as allophones.*

   The L2 learner's L1 displays allophonic variation where the phoneme /n/ is realized as [ŋ] if it is followed by a velar consonant, as is for instance the case in Italian. In Norwegian a phonemic opposition between the two nasals must be learnt.

4) *One may add a fourth category: only one allophone of an allophonic pair exists in the native language.*

   The context-dependent (stress, preceding sound) contrast between [±aspirated] allophones of /p/, that is [p] – [pʰ], is in general difficult for speakers with only non-aspirated realizations in their L1. Lack of aspiration in positions where it is required leads to confusion for Norwegian L1 speakers as aspiration is a main contributor to the opposition between Norwegian /p/ and /b/.

3.2. Level of phonetic detail

The choice of exercises for oppositions depends on the level of detail represented in the database. Since L1-L2*map* makes use

of the multi-lingual UPSID database, its granularity is determined by the features used in that database. UPSID contains the phoneme inventories of 451 languages. The phonological features that are used can (at least) distinguish any phoneme opposition in any of the languages.

The representation of phonemes in a language directly affects the output of a contrastive analysis. In the UPSID database, Mandarin Chinese plosives /t,d/ are classified as dental/alveolar, without a more precise evaluation of the place of articulation. But [12] classifies Mandarin Chinese plosives /t,d/ as alveolar, which makes them different from Norwegian /t,d/, which according to [13] are dental. The output of the analysis will consequently generate exercises for Chinese-speaking students where the Norwegian dental articulation is focused. By choosing another reference for Chinese, for example [14], there will be no exercises generated for the Chinese NSL learners, as the Mandarin Chinese plosives are classified as dental. The choice of the reference on which to base the phoneme inventory description is of crucial importance for the resulting contrastive analysis, and therefore for the exercises which are suggested to the language learner.

L1-L2map, and the effect of phonetic detail on the pronunciation target in a CAPT system, is described in greater detail in [15]. The article also discusses several reasons for introducing consonant phoneme inventories for each syllable position (onset, nucleus, coda) in L1-L2*map*. Positional information is shown in an efficient and intuitive manner in the L1-L2*map* tabs. The article also suggests several useful extensions of the tool. We refer to the article for a further motivation of the level of detail adopted in L1-L2*map*.

3.3. Embedding L1-L2*map* in Ville-N

To be able to carry out a contrastive analysis for a specific learner of Norwegian, he/she is asked to supply information about his/her native language, and select one of the Norwegian dialects as the target dialect when logging in for the first time. On the basis of this information, the Norwegian CAPT system sends a request to L1-L2map, which is implemented as a client-server system. The system returns a list of the sounds in Norwegian which are not part of the inventory of the learner's native language. The list is in xml- or json-format, and can be linked to exercises in the CAPT system. The exercises can be selected by the user from a pull-down menu in Ville-N – comparable to the selection of the category "Pronomen" (E. "pronoun") in Figure 1.

It should be pointed out that a phonetically naive user does not see the contrastive analysis, which is performed automatically in the background. For users who are interested, L1-L2*map* is accessible from Ville-N.

3.4. Contributing to L1-L2*map*

L1-L2*map* is a tool for contrastive analysis that can be used for any language. It is possible to edit or add languages using a very simple and intuitive interface, and the tool can then be used to generate a list of unfamiliar, and therefore possibly problematical segments for learners of the selected target language, given any source language in the database.

L1-L2*map* is a platform which allows language experts to make language data available in a format that can be used directly by technologists building CAPT systems. Although any user has access to L1-L2*map*, only a group of language experts can insert data about their native language and/or other languages that they have near-native phonetic-linguistic competence for. An instruction for the use of L1-L2*map* is given on the webpage [6]. Language experts who wish to have editing rights can send an e-mail to l1-l2map@ntnu.no.

4. Conclusions

The most salient characteristic of Norwegian from a language learner's point of view is that it does not have an accepted standard pronunciation. For effective communication, learners must familiarize themselves with the different dialects, which are widely used − unlike in other languages, where native speakers can usually adapt their speech towards an accepted norm when they interact with foreigners. The Norwegian CAPT system Ville-N therefore comprises 8 different talking heads: one male and one female tutor for each of four dialects (Section 2). To practise their pronunciation, learners can select one of the tutors as a role model.

To guide language learners only to relevant exercises, a contrastive analysis of the NSL learner's native and target language (dialect) is performed by L1-L2*map*. This tool returns a list of sounds which may cause problems for the learner. The list can be used to select relevant listening and speaking exercises. L1-L2*map* contains over 500 languages, and can be used by other CAPT researchers (Section 3). This paper therefore ends with an offer to use the tool, which is available on a wiki, and an invitation to contribute to the language descriptions.

5. Acknowledgements

6. References

[1] Wik, P. and Hjalmarsson, A., "Embodied conversational agents in computer assisted language learning", Speech Communication 51(10), 1024-1037, 2009.

[2] "Common European Framework of Reference for Languages: learning, teaching, assessment" (p. 110), online: http://www.coe.int/T/DG4/Portfolio/documents/Framework_EN.pdf, accessed on 18.04.2011.

[3] UVic Humanities Computing and Media Centre and Half-Baked Software, "UVic's Language Teaching Clipart Library", online: http://hcmc.uvic.ca/clipart/, accessed on 18.04.2011.

[4] Husby, O., Høyte, T., Nefzaoui, S.J., Nordli, I., Robbins, S. and Øvregaard, Å, "An Introduction to Norwegian dialects", Trondheim: Tapir, 2008.

[5] McAllister, R., "Second language perception and the concept of foreign accent", Speech Techn. in Lang. Learning, 1998.

[6] Bech, Ø., Koreman, J., Husby, O. and Wik, P., "L1-L2*map*", online: http://calst.hf.ntnu.no/l1-l2map, accessed 18.04.2011.

[7] Maddieson, I. "Patterns of Sounds", Cambridge: CUP, 1984.

[8] "Handbook of the International Phonetic Association. A Guide to the Use of the International Phonetic Alphabet", Cambridge: CUP, 1999. Online: www.langsci.ucl.ac.uk/ipa/ipachart.html, accessed 18.04.2011.

[9] Lado, R., "Linguistics across Cultures: Applied Linguistics for Language Teachers", Ann Arbor: Univ. of Michigan Press, 1957.

[10] Avery, P. and Ehrlich, S., "Teaching American English pronunciation", Oxford: Oxford University Press, 1992, p. xvi.

[11] Eckman, F., Elreyes, A. and Iverson, G., "Some principles of second language phonology", Second Language Research 19, 169-208, 2003.

[12] Sun, C., "Chinese: A Linguistic Introduction", Cambridge: CUP, 2006.

[13] Vanvik, A., "Kort innføring i fonetikk", Oslo: Universitets-forlaget, 1983.

[14] Duanmu, S., "Phonology of Chinese (Mandarin)", in: Encyclopedia of Language and Linguistics, 2nd edition, Elsevier Publishing House, 2005. Online: http://www-personal.umich.edu/~duanmu/ELL05.pdf, accessed 15.05.2011.

[15] Koreman, J., Bech, Ø., Husby, O. and Wik, P. (submitted). "L1-L2*map*: a tool for multi-lingual contrastive analysis", ICPhS.