Contrastive analysis through L1-L2map

Preben Wik[1], Olaf Husby[2], Åsta Øvregaard[2], Øyvind Bech[2], Egil Albertsen[2], Sissel Nefzaoui[2], Eli Skarpnes[2], Jacques Koreman[2]
[1] *Department of Speech, Music and Hearing, KTH, Stockholm, Sweden*
[2] *Department of Language and Communication Studies, NTNU, Trondheim, Norway*

Abstract

This paper describes the CALST project, in which the primary aim is to develop Ville-N, a computer assisted pronunciation training (CAPT) system for learners of Norwegian as a second language. Ville-N makes use of L1-L2map, a tool for multilingual contrastive analysis, to generate a list of language-specific features. These can be used to tailor pronunciation and listening exercises. The tool can also be used for other target languages.

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 Royal Institute of Technology (KTH), the University of Oslo (UiO), The Adult Education Centre (EVO) in Trondheim, and. The project's aim is to develop a computer-assisted pronunciation training (CAPT) system for Norwegian as a second language (NSL). The project has two main goals:

1) To develop a Norwegian CAPT system, Ville-N, based on Ville, the virtual language teacher for Swedish, developed at KTH (Wik, 2004). The system should be able to cater for L2-learners with a wide range of communicative abilities, ranging from foreign university students at NTNU and UiO who are proficient in English, to illiterate users at EVO with no English skills. The program should be used with minimal requirements on instruction or supervision.

2) To create a database and tools for contrastive phonetic and phonological analysis for all relevant L1-L2 pairs, partly as a foundation for future research in second language acquisition (SLA), and partly in order to better tailor exercises used in Ville-N.

Ville-N will be used to complement pronunciation teaching in the Norwegian courses both for foreign students and employees at NTNU and UiO, 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 also include L2-learners that are illiterate and from other language and social backgrounds than what is found at the universities.

Several extensions to the Swedish Ville have been made in order to accommodate the different needs of the Norwegian system.

Norwegian dialects

The Norwegian language situation is quite different from that of Sweden and many other language communities in that there is no accepted pronunciation standard in Norwegian. Although there is a common form taught in adult second language classrooms called Bokmål (Urban East Norwegian, UEN), different dialects are used both in formal and informal situations. This creates a serious problem for L2-learners of Norwegian since there are large pronunciation differences in the various dialects and often different words are used to express the same meaning. This dialectal variation cannot easily be addressed in standard language courses, and a need has thus been identified to better equip L2-learners to deal with everyday communicative situations where variation in the speakers' dialect is typical

Multiple speakers

To address this problem recordings of multiple speakers have been made. One male and one female speaker of the dialects in the following regions of Norway have been recorded: Northern Norway, Trøndelag, Western/Southern

Norway and Southeastern Norway (UEN). All in all 8 different voices will be used in the program, accompanied by 8 different embodied conversational agents (ECAs), to give each voice a personality.

The learner can select an ECA from the GUI and listen to and practice one specific dialect, or train across dialects in the same exercise and let the program select the target voice.

As reported by several researchers, including McAllister (1998), it is good to listen to many different speakers in order to achieve better listening comprehension. Multiple-talker models have been reported as particularly effective to improve perception of novel contrasts (c.f. Logan et al., 1991; Probst et al., 2002), as the inherent variability allows for induction of general phonetic categories or other L2 specific salient features.

Learners will also be able to run the exercises with one specific dialect in mind, which is useful when selecting a role model for pronunciation in production exercises. The advantage of having both male and female voices in each dialect becomes apparent in this case to allow learners to choose a role model with the same gender as themselves.

Hyper-articulated recordings

The native speakers who made the recordings were informed to speak as they normally do with normal speed, reductions and coarticulations, in order to offer the learners spoken utterances that are as close to authentic speech as possible. It is however difficult for beginners to assimilate reductions and coarticulations in the early phases of learning, and it is often common practice to speak slower and clearer when speaking to L2 learners who are in an early phase of development. To incorporate such considerations in the program two options were considered: To make it possible for the learners within the program to slow down the speech samples by manipulating the acoustic signal, or to make double recordings, one normal and one slow hyper-articulated version of all the recordings. The difference between slow and normal speech is not uniformly distributed, as for example plosives are not stretched in the release burst but only in the occlusion. Long vowels are typically exaggerated, coarticulations will be reduced or removed, and other aspects of the speech such as the lexical stress will be affected differently with

stressed syllables being stronger and more emphasized in a hyper-articulated version. It was hence decided that to ensure the best possible quality in the learning material, exaggerated, hyper-articulated versions of all recordings were made even if it would include more work and a larger set of recordings.

The learners have the option to choose to do the exercises using either type of recording, or a learner can choose to use normal recordings and select a: "Say again" button, to get a repetition of the last spoken word in a hyper-articulated version.

Wordlists

A wordlist of basic vocabulary from the course books used by the participating institutions has been created and categorized into semantic categories. Approximately 1000 words were selected from the aggregated wordlists and divided into 43 categories. The criterion for the selection of the base vocabulary was also that it should satisfy the A1 and A2 vocabulary range of CEFR, (Common European Framework of Reference for Languages)

All words have then been visualized. Approximately 30% of the images are from "UVic's Language Teaching Clipart Library". For words where no appropriate image was found complementary drawings have been made by a local artist in the same artistic style as the drawings from the internet, in order to get a consistent and coherent set of images.

English translations, transcription and the inflection of words have been added, and as mentioned above, sound files have been recorded for the aggregated wordlist, in four dialects, with one male and one female speaker for each dialect, and in both normal and hyper-articulated versions.

The L1-L2map

The contrastive analysis hypothesis (CAH) as presented in Lado (1957), claims all problems in learning a foreign language can be explained from transfer problems induced by the learner's native language. It is today generally accepted that the claims made by the CAH are too strong, and that there are other factors which determine the difficulty language learners have with acquiring new sounds (Eckman, 1977; Odlin, 1989; Flege, 1995; Major, 2001). This does however not mean that CAH should be

completely rejected or abandoned. As stated by Ellis (1994)

"*The problem with CAH is that it is too simplistic and too restrictive. The solution as many researchers have come to recognize, lies not in its abandonment but in careful extension and revision*"

The second task in the CALST project is to design and evaluate a revised and extended contrastive analysis tool called L1-L2map. Since the problematic aspects of the L2 are not the same for all learners, the aim of the CAPT system is to assist learners (and teachers) in identifying the problematic aspects for each individual, and work on these contrasts with special exercises. The idea with L1-L2map is to make the CAPT system informed by L1 specific filtering.

The L1-L2map will serve as a platform for researchers with a phonetic background to encode language data and make it available in a format that can be used by CAPT creators. L1-L2map is designed as a wiki with two levels of users.

First of all as a generally accessible tool, where any user can access the data and browse and compare the phonological features of different languages. A group of specialists will have administrative rights and the responsibility for inserting feature data about languages that they have phonetic-linguistic competence for. Each language is thus encoded individually by an expert in that particular language, and the contrastive analysis is performed by 'superimposing' the data for two languages on top of each other.

The L1-L2map can 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. The first version of the L1-L2map is based on the UPSID database, which contains 451 languages (Maddieson, 1980). This number has been increased to more than 500 languages in L1-L2map. As shown in Figure 1, the result of the contrastive analysis is displayed in four tabs.

The choice of tabs is based on the IPA representation: "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 IPA was used in order to provide language experts using the system with a simple and recognizable lay-out.

As shown in Figure 1, a color scheme is employed in the visualization of the data. The first language chosen is blue and assumed L1 status, the second language chosen is red and assumed L2 status, and the features they have in common (where there is an overlap) are green. This way it becomes apparent for the L2 learner which features are different from their L1 (red), and thus needs attention since they are absent in the L1.

Extensions to UPSID

The UPSID database only lists sounds that are distinct phonemes in any given language. It is not enough to do a simple comparison of which sounds constitute phonemes in each language. A number of extensions are in the process of being added to the L1L2-Map. These include positional restrictions, syllable structure (phonotactics), tone, stress, and timing.

Positional phonemic restrictions

The positions in which sounds occur in syllables must be taken into account. Difficulties that could be predicted from CAH if position is part of the description, will otherwise go by unnoticed. For example in Mandarin, only two consonants ([ŋ] and [n]) are allowed at the end of a syllable, even if many other consonants appear in syllable initial position. A consequence of this is that many consonants which can occur at the end of Norwegian syllables present a difficulty for learners of Norwegian with Mandarin as L1.

The L1L2-map presented in Figure 1 is the complete set of consonants in both Norwegian and Mandarin, taken from the UPSID data, where positional restrictions are not taken into account. In the next version of L1-L2map, this pane will be divided into three separate panes, with each pane only displaying the phonemes that are allowed in initial medial, and final position respectively.

| | Languages ▼ | Sounds ▼ | | | Mandarin | Norwegian | | Username: | Password: | Log in |

| | Consonants | Consonants (other) | Vowels | Diphthongs | Language information |

	Labial		Coronal				Dorsal			Radical		Laryngeal
	Bilabial	Labio-dental	Dental	Alveolar	Palato-alveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Epi-glottal	Glottal
Plosive	pʰ p b		tʰ ɖ	*tʰ *t		tʰ ɖ		kʰ k g				
Nasal	m		n̪	n		ɳ		ŋ				
Trill												
Tap, Flap				ɾ								
Fricative		f	s	ʃ	ʂ ʐ ç				χ			h
Lateral fricative												
Approximant		ʋ					j ɥ	w				
Lateral approximant				*l l		ɭ						
Lateral flap												

Figure 1 Visualization of the consonant part of the L1-L2map. Here exemplified with Mandarin as L1 in blue, Norwegian as L2 in red. Overlapping phonemes displayed in green.

Syllable structure (phonotactics)

Not only positional restrictions, but also the phonotactic constraints of languages (restrictions on the permissible combinations of phonemes) will be encoded in the L1L2-map. Norwegian has a relatively complex syllable structure comparable to that of Swedish, and constitute a difficult part of acquiring the language for many learners.

Permissible consonants and consonant clusters in onset and coda will be encoded into separate lists in the database, and when a contrastive analysis is performed between the learners L1 and the L2 clusters that are allowable combinations in the L2 but missing from the learners L1 will be extracted.

L1-L2map is a useful tool for developers of CAPT systems for any language, as well as for language teachers. The tool allows the insertion of new languages and/or dialects, and outputs useful information about the L2 phones which L2 learners need to acquire.

The L1-L2map is available online: http://calst.hf.ntnu.no/l1-l2map

Acknowledgements

References

Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning, 27*(2), 315-330.

Ellis, R. (1994). *The Study of Second Language Acquisition.* Oxford University Press.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, 233-277.

Lado, R. (1957). *Linguistics across cultures.* University of Michigan Press.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English/r/and/l: A first report. *Journal of the Acoustical society of America, 89*(2), 874-886.

Maddieson, I. (1980). *UPSID: UCLA phonological segment inventory database.* Phonetics Laboratory, Department of Linguistics.

Major, R. C. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology.* Lawrence Erlbaum.

McAllister, R. (1998). Second language perception and the concept of foreign accent. In *STiLL-Speech Technology in Language Learning.*

Odlin, T. (1989). *Language Transfer.* Cambridge University Press.

Probst, K., Ke, Y., & Eskenazi, M. (2002). Enhancing foreign language tutors-In search of the golden speaker. *Speech Communication, 37*(3-4), 161-173.

Wik, P. (2004). Designing a virtual language tutor. In *Proc of The XVIIth Swedish Phonetics Conference, Fonetik 2004* (pp. 136-139). Stockholm University.