EXPANDING CALST: MULTILINGUAL ANALYSIS OF L1-L2 PHONOTACTICS FOR LANGUAGE TEACHING

Violeta Martínez-Paricio, Jacques Koreman, Olaf Husby, Jardar Eggesbø Abrahamsen, Øyvind Bech, Norwegian University of Science and Technology

The present article describes the current expansions of L1-L2*map* and CALST, two platforms for L2 pronunciation teaching developed at the Norwegian University of Science and Technology (NTNU). L1-L2*map* is a multi-lingual database that allows comparison of the phoneme inventory of a large number of languages. The result from contrastive analysis in L1-L2*map* is used in CALST, a Norwegian computer-assisted pronunciation training (CAPT) system, where pronunciation exercises are selected depending on the native language of the user. Originally, L1-L2*map* and CALST exclusively contained segmental information and exercises for practicing the perception and production of individual sounds. In this paper we show that both tools are currently being expanded to include prosodic information too. In particular, we discuss the possibility of incorporating multilingual data on consonant clusters and positional restrictions of segments. This will enable learners to practice their perception, production and spelling of not only individual sounds, but also the specific positional and combinatorial restrictions associated with them in a given language. Future expansions of L1-L2*map* and CALST will also incorporate information and exercises on other prosodic properties such as stress, tone and intonational patterns.

INTRODUCTION

In second language courses, it is common to encounter learners with very different language backgrounds, especially when the language course takes place in the country where the target language is spoken. For instance, learners with varying native languages participate together in the same Norwegian courses taught at the Norwegian University of Science and Technology (NTNU) in Trondheim. Among others, we can find Vietnamese, Spanish, Italian, German, Polish, Thai, French, Russian and Mandarin Chinese speakers attending the same Norwegian course. In such a teaching context, or similar ones, it is difficult to train students in the pronunciation of a second language (L2): the difficulties learners experience often vary depending on their native language (L1). Furthermore, teachers do not normally have an in-depth knowledge of the phonetics of all course participants' native languages. Therefore, they do not always know beforehand which are the most problematic sounds for a particular speaker. Finally, there is usually not enough time in class to give corrective feedback to individual language learners (Koreman *et al.* 2011:1). To solve all these challenges and help both L2 teachers and L2 students, we have developed two platforms at NTNU to train the pronunciation of an L2, which takes into account the differences between the L1 of the learner and the L2 she aims to learn.

First, we created L1-L2*map* (http://calst.no/L1-L2map), a multilingual database which allows comparison of the phonemic inventory of two languages: a potential L1 and a potential L2. The segmental information in this database is presented in very basic phonetic charts which are based on the International Phonetic Alphabet charts, and use a simple color-coding to display the similarities and differences between two languages. We indicate in green the segments that are

common to the two languages, in blue the segments that only occur in the L1 of the speaker and, in red, the segments that only occur in the L2. The sounds in red are the most important from the teaching perspective, since speakers tend to experience greater problems with the production and distinction of unfamiliar sounds (i.e. sounds absent in their native languages). This information tells teachers which are the most problematic sounds for each learner, enabling them to put greater effort in correcting their production/perception. However, as has just been acknowledged, there is often not much time in class to practice and correct one by one all students, especially when they come from different language backgrounds.

To solve these problems, we have created a second program: CALST, the *Computer Assisted Listening and Speaking Tutor*, a computer-assisted pronunciation training (CAPT) system which allows students to practise their pronunciation, listening and spelling skills for the main Norwegian dialects, given there is no accepted pronunciation standard in Norwegian. CALST is is based on VILLE (Wik 2011) and it is available online (http://calst.no). Learners can use it for free, not only during class hours, but whenever they have time to work on their Norwegian pronunciation skills. Importantly, CALST benefits from the contrastive analysis carried out in L1-L2*map*. The first time the learner logs in CALST, she must indicate her native language. Once this information is registered, the program automatically links the sounds in red in L1-L2*map* (i.e. the problematic sounds) to specific exercises in CALST. In this way, each learner gets to practise the unfamiliar sound contrasts in the L2 depending on her specific L1. In the case at hand, the listening and pronunciation exercises have been developed and recorded only for several Norwegian dialects. Hence, learners of Norwegian can already benefit from this tool. However, it is very important to highlight that CALST is modular: once exercises are developed for other languages (e.g. English, Spanish, French, etc.), they can be easily integrated in the system for learners of that language. In that sense, learners of other languages will also be able to benefit from the multilingual contrastive analysis carried out at L1-L2*map*. In addition to the exercises for new sounds, all learners have access to some general vocabulary exercises in CALST. In these exercises, learners start familiarizing themselves with the sounds of the L2. The vocabulary exercises allow practising listening, pronunciation and spelling skills with approximately a thousand words and basic expressions.

Current expansions: prosody in L1-L2*map* and CALST

The two platforms, L1-L2*map* and CALST, originally arose to train learners on the acquisition of new sound contrasts and individual segments not present in their native languages. However, as pointed out by Young-Scholten and Archibald (2000: 64), "it is not just the ability (or lack thereof) to produce an individual segment that results in second language accent; a second language learner must also be able to combine the segments in the sequences demanded by the target language". This ability of combining segments and/or being able to produce them in a particular position is not a skill that needs to be practised only to avoid second language accent, but more importantly, the learner needs to train this skill to ensure her intelligibility. Consider, for example, a Vietnamese learner of Norwegian. Vietnamese speakers do not have any trouble in pronouncing Norwegian words with the sound that corresponds to the phoneme /f/ in an onset position, since Vietnamese allows /f/ in such position. However, Vietnamese /f/ never occurs in coda position. When Vietnamese speakers have to pronounce Norwegian words with [f] in coda position, they sometimes struggle. Very often, instead of realizing a coda [f], speakers tend to substitute it for a labial voiceless unreleased stop [p̚], especially at a beginner level. From a

communicative perspective, this can be very problematic since Norwegian speakers may find it difficult to understand what the Vietnamese speaker is trying to say, since the pronounced word may have a different meaning.

Another example of the importance of learning more than just the individual segments of the L2 is found in the interlanguage of speakers whose native languages have very simple (or relatively simple) syllable structures, but the L2 they want to master contains complex syllables. Speakers of languages with simple syllables generally have difficulties with the pronunciation of complex onsets and complex codas. For instance, even if Spanish speakers are able to pronounce /s/ and /k/ in a simple onset position, when the two consonants occur in a complex onset in the L2, as in English ***sk****ype*, Spanish learners of English tend to repair this complex onset, absent in their native language, by inserting a vowel before such complex onset. Therefore, instead of producing the monosyllable *skype* they often say something closer to [**e**s.kajp] or [**ə**s.kajp], with two syllables (we use the dots to indicate syllable boundaries). In a similar way, Thai speakers often repair similar complex onsets in the L2 and in loanwords by inserting a reduced vowel after the s-, splitting the complex onset, e.g. [sᵊ.kaj]. (Kenstowicz & Suchato 2006). The example also shows that, due to a native cooccurrence restriction in Thai, the final /p/ is often deleted after a glide in the interlanguage of these speakers (even though /p/ can occur syllable-finally in the language). In sum, "while factors independent of L1 seem clearly to be a factor in the acquisition of L2 syllable structure and phonotactics, native language restrictions clearly affect not only the production but also the perception of L2" (Broselow & Kang 2013: 540). For all these reasons, we believe it is crucial to incorporate this type of phonotactic information in our multilingual database, as well as in the Norwegian version of CALST, so learners of an L2 also get to practise the segments in varying combinatorial and positional contexts.

In the remainder of the paper we first present the methodology we followed for incorporating the multilingual data on consonant clusters, illustrating how this complex information can be visualized in L1-L2*map*. Then, we discuss the type of listening and pronunciation exercises we are currently developing so that learners of Norwegian can practice the different consonant sequences allowed in the language. The article ends by discussing the possibility of expanding L1-L2*map* and CALST even more, so that both systems incorporate in the future additional prosodic information, such as multilingual data related to the position of stress and to lexical tone as well as basic intonational patterns in different languages.

CONSONANT CLUSTERS IN L1-L2*map*

L1-L2*map* allows comparing the phonemic inventory of any two languages from a set of around 500 languages. The segmental data displayed at L1-L2*map* are mainly extracted from UPSID (Maddieson 1984) and LAPSyD (Maddieson et al. 2011), two databases available online which display the phonemic inventory as well as some basic phonological information about several hundred languages. Additional segmental information of other languages has been added to L1-L2*map*. Our comparative database is implemented as a wiki so that language experts can include new data about a particular language and/or correct the existing data. The current goal of our project is to expand L1-L2*map* so that it can present and compare information about the

phonotactics of languages as well, i.e. information about the specific positional and combinatorial restrictions of segments in a given language.

To be able to deduce distributional differences from a contrastive analysis, a complete list of all the possible onsets and codas for the two languages, the L1 and the L2, is required. Unfortunately, UPSID and LAPSyD do not contain very detailed information regarding the phonotactic and syllabic restrictions active in languages. Consequently, we cannot easily extract and compare this information in L1-L2*map*. We do find information in LAPSyD regarding the maximum/minimum size of syllables in a given language, indicated in the form of syllable templates with a simple notation employing <V> for vowels, <C> for consonants, and brackets for optional material. For instance, a language that (i) does not allow onsetless syllables, (ii) bans complex codas and (iii) allows maximally two consonants in complex onsets is coded as having the following syllable template: (C)CV(C). This information is valuable because it can be inferred that speakers of such a language might have problems when acquiring an L2 with more complex codas and onsetless syllables. Very often, speakers tend to adjust unfamiliar structures in the L2 to match the restrictions in their L1. Therefore, to avoid the production of complex codas in the L2, speakers of this language might delete one or more segments in the coda (e.g. CV$\mathbf{C_1C_2}$ > CV$\mathbf{C_1}$). Alternatively, they can insert a vowel to split the complex coda (e.g. CV$\mathbf{C_1C_2}$ > CV.$\mathbf{C_1\underline{V}C_2}$). Additionally, to ensure that a syllable always has an onset, a speaker of this L1 might insert a glottal stop (or another consonant) in an onsetless syllable in the L2.

If we only compare these types of syllabic templates in an L1 and an L2, however, we will probably get an oversimplified impression of which are the most problematic areas of pronunciation for a particular L2 learner. For example, it might be the case that the language with the (C)CV(C) syllable template only allows nasal codas. In that case, it is very likely that when acquiring an L2 that also permits obstruents and liquids in the coda, the learner will need to pursue some training on the pronunciation of such segments in that specific syllable-final position. Furthermore, it can be the case that this same language only allows complex onsets that follow the Sonority Sequencing Principle (Selkirk 1984, Clements 1990), i.e. complex onsets that rise in sonority towards the nucleus of the syllable (e.g. [pl-, sn-, tj-]). The speakers of this L1 will also need to practice the pronunciation of other types of complex onsets, such as the ones composed by the fricative /s/ and a voiceless stop. These complex onsets (e.g. [sp-, st-, sk-]) do not follow the Sonority Sequencing Principle, but they are frequently attested in many languages from different linguistic families. Assuming that a speaker of an L1 which allows [pl-, kl-] will not have any trouble in pronouncing complex onsets like [sp-, sk-] is probably wrong. That is exactly why we need a more complete database which allows us comparing not only the maximum and minimum size of syllables across languages, but their actual segmental make-up.

To pursue this goal in L1-L2*map*, we started recompiling specific syllabic information for the following ten languages: Norwegian, Spanish, English, Mandarin Chinese, Farsi, Vietnamese, Thai, Polish, Tagalog and Japanese. These languages were selected for both practical and typological reasons. On the one hand, we were interested in having data for languages representative of various language families, with different syllabic restrictions. On the other hand, speakers of such languages are quite frequent in Norwegian courses and, therefore, having this information can already be of help for learners and teachers of Norwegian, as well as for the

development of the Norwegian version of CALST. We hope that L1-L2*map* will be expanded with additional data in the future in collaboration with other languages experts.

We first coded the segments and combination of segments that can occur *before* the syllable sonority peak (i.e. which is generally a vowel, but in can also be a consonant in languages that allow syllabic consonants) in the above-mentioned languages. Secondly, we coded the information related to the specific segments and combinations of segments that can occur *after* the syllable sonority peak. We followed this methodology, distinguishing pre-peak and post-peak material, to avoid analytical problems related to the particular association of prevocalic and postvocalic glides, which in some languages are analyzed as being part of the nucleus, but in others as being part of the onset for rising diphthongs or of the coda for falling diphthongs (see Smith 2007 for discussion). For teaching purposes, it is crucial to know whether both the L1 and the L2 admit, for example, a sequence like [Liquid+Glide] at the beginning of a syllable, before the syllable peak. If linguists analyze the glide in L1 as being part of the onset, but in the L2 as part of the nucleus (together with the syllable peak) it is in principle not relevant for teaching the pronunciation of the L2. The relevant fact is that the two languages are *equal* with respect to the presence of the sequence [Liquid+Glide]. When the speaker gets to practise the realization of complex sequences in the L2 in CALST, she will not get exercises focused on the sequence [Liquid+Glide]. Only speakers with an L1 that lacks the [Liquid+Glide] sequence at the beginning of a syllable would get such exercises.

By providing only the possible pre-peak and post-peak sequences of languages, we miss some important linguistic generalizations since, very often, there are restrictions that hold over the whole rhyme (i.e. the nucleus and the coda) or over the onset and the nucleus. For instance, In Urban East Norwegian the vowels [e, ɛ] are in general blocked in front of the consonants /r/ and /ɽ/ and also before the glides /j/ and /w/ (Kristoffersen 2000:14). However, annotating all the possible syllables of a language –rather than just the possible pre-/post-peak sequences– would substantially increase the complexity of each language description, making the comparison between languages very difficult and redundant. Furthermore, we believe the L2 learner will indirectly learn this phonological, structural information when she is exposed to specific data and real examples in CALST. What is important for the learner is to familiarize herself with the perception and production of unusual consonants clusters.

We illustrate below how the consonant cluster information is presented in L1-L2*map*. Figure 1 displays the lay-out of a comparison of the pre-peak consonant sequences of two segments in Norwegian (L1) and English (L2). The color coding we employ here is exactly the same as in the segmental comparisons: (i) blue is used to marked complex CC- clusters present only in the native language of the learner (i.e. Norwegian), (ii) red indicates CC- clusters permitted only in the L2 (i.e. English) and (iii) green is used to signal CC- clusters which are common to both languages. Crucially, the sequences highlighted in red (i.e. the unfamiliar sequences for the learner) will be linked to specific production and perception exercises in CALST, to train the learners in the acquisition of these new pre-peak consonant clusters.

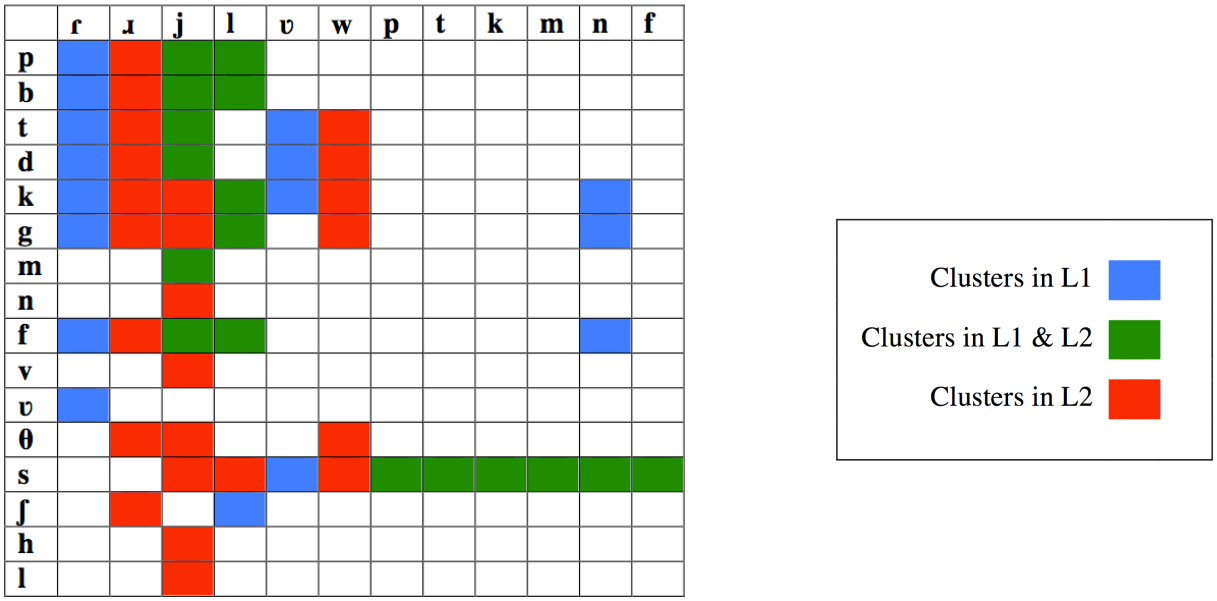

Figure 1. Lay-out of the result of a contrastive analysis comparing prevocalic consonant clusters of two segments (CC-) in (southeastern) Norwegian (L1) with English (L2)

CONSONANT CLUSTERS EXERCISES IN CALST

The original exercises in CALST consist of (i) listening (Listen & Click), (ii) pronunciation (Listen & Speak), (iii) spelling (Listen & Write). In these exercises the learner gets familiarized with general vocabulary, but also with new sound contrasts and individual segments (see Koreman et al. 2013 for a detailed explanation of these exercises). We are now working on the expansion of the same exercises so that the learner can also practice unfamiliar consonant sequences allowed in Norwegian. Consonant combinations highlighted in red in the contrastive analysis of consonant clusters in L1-L2*map* are linked automatically to exercises containing those sequences in Norwegian, so that learners can practise these unfamiliar sound combinations. The learners are presented with the exercises and do not normally see the information in L1-L2*map,* which is intended for phonetically trained developers and teachers. In the listening exercises (Listen & Click), a series of similar words is presented visually to the learner, where one of the words contains an onset or coda cluster, while the rest of the words have undergone some of the most common repair strategies attested in L2 interlanguage, such as deletion of a segment, metathesis, or feature change. For instance, the learner sees the Norwegian words <sta> - <da> - <ta> - <sa> (E. “stubborn”, “there”, “take” and “said”, respectively) on the computer screen, while one of the words is pronounced by the virtual tutor. The learner has to choose the word she is hearing by clicking on it on the computer screen. In pronunciation exercises (Listen & Speak), learners train the pronunciation of words with unfamiliar onsets or codas. After listening to the artificial tutor’s pronunciation of the word, the user records herself and plays the recording of her own voice followed by the realization of the same word pronounced by the artificial tutor. By comparing these, she can self-evaluate her production. In future, we aim to incorporate automatic speech recognition techniques in CALST, so that the artificial tutor will be

able to detect pronunciation errors. Automatic speech recognition will also help to detect vowel epenthesis, which cannot easily be addressed in the exercises presently available in CALST. Finally, spelling exercises (Listen & Write) will further strengthen the learner's association of the perceived clusters with its orthographic representation.

FUTURE EXPANSIONS AND CONCLUSIONS

L1-L2*map* and CALST arose to train L2 learners in the acquisition of new segments and new sound contrasts. We have explained in this paper how both platforms are currently being expanded to incorporate information about consonant sequences too. Still, the acquisition of an L2 sound system does not only imply learning the new segments and their phonotactic restrictions in the L2, but leaners also must acquire a new prosodic system. That is, learners must acquire other linguistic properties in the L2 related to: (i) the position and phonetic cues of stress (in cases where the L2 has stress), (ii) the tonal patterns (in languages with lexical tones, but also in pitch-accent languages) and (iii) the intonational patterns of the language (Broselow & Kang 2013).

Once we have developed the exercises for training the production and perception of particular consonant clusters, we hope to extend L1-L2*map* and CALST with information about the position of stress, making use of the StressType (http://st2.ullet.net/) database, which contains information about stress in a large number of languages.

Finally, as we indicated in the introduction of this article, we hope to develop CALST with exercises for other languages so that leaners of other L2s can also benefit from this multilingual approach to language teaching.

ACKNOWLEDGMENTS

We gratefully acknowledge the financial support from VOX, the Norwegian Agency for Lifelong Learning, as well as the Ministry of Children, Equality and Social Inclusion and the Ministry of Local Government and Modernisation and the Ministry of Labour and Social Affairs.

ABOUT THE AUTHORS

Violeta Martínez-Paricio (prosodic typology), Jacques Koreman (phonetics), Olaf Husby (Norwegian as a Second Language), Jardar Eggesbø Abrahamsen (phonology) and Øyvind Bech (developer) collaborate in the cross-disciplinary CALST project (Computer-Assisted Listening and Speaking Tutor) at the Department of Language and Literature at the Norwegian University of Science and Technology (NTNU) in Trondheim, Norway. The project employs a contrastive analysis approach to language learning. The research group has developed and implemented pronunciation training exercises for Norwegian, but their multi-lingual platform for language learning makes the same exercises immediately available for other languages when learning content for other languages is inserted. The platform intends to make pronunciation training available both in and outside the classroom.

REFERENCES

Broselow, E. & Kang, Y. (2013). Phonology and Speech. In J. Herschensohn & M. Young-Scholten (eds). *The Cambridge Handbook of Second Language Acquisition* (pp. 529–544). Cambridge: Cambridge University Press.

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston and M. E. Beckman (eds). *Papers in Laboratory Phonology I: Between the grammar and the physics of speech* (pp. 283–333). Cambridge: Cambridge University Press.

Kenstowicz, M. & Suchato, A. (2006). Issues in loanword adaptation: A case study from Thai. *Lingua*, *116* (7), 921–949.

Kristoffersen, G. (2000). *The phonology of Norwegian*. Oxford: Oxford University Press.

Koreman, J., Bech, Ø., Husby, O. & Wik, P. (2011). L1-L2*map*: a tool for multi-lingual contrastive analysis. *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS2011),* Hong Kong.

Koreman, J., Wik, P., Husby, O. & Albertsen, E. (2013). Universal contrastive analysis as a learning principle in CAPT, *Proceedings of the workshop on Speech and Language Technology in Education (SLaTE 2013),* ISSN 2311-4975, pp. 172-177.

Maddieson, I. (1984). *Patterns of Sounds*. Cambridge: Cambridge University Press

Maddieson I., Flavier S., Marsico E., Pellegrino F. (2011). LAPSyD: Lyon-Albuquerque Phonological Systems Databases, Version 1.0. http://www.lapsyd.ddl.ish-lyon.cnrs.fr/

Selkirk, E. (1984). On the major class features and syllable theory. In M. Aronoff & R. T. Oehrle (eds). *Language Sound Structure: Studies in Phonology* (pp. 107–136). Cambridge: MIT Press.

Smith, J. L. (2007). Representational complexity in syllable structure and its consequences for Gen and Con. In S. Blaho, P. Bye & M. Krämer (eds). *Freedom of analysis?* (pp. 257–280). Berlin: Walter de Gruyter.

Wik, P. (2011). The Virtual Language Teacher. Ph.D. thesis, KTH School of Computer Science and Communication, 2011.

Young-Scholten, M. and Archibald, J. (2000). Second Language Syllable Structure. In J. Archibald (ed). *Second Language Acquisition and Linguistic Theory* (pp. 64–101). Oxford: Blackwell Publishers.