A systematic approach to the pronunciation training of phonotactics

Jacques Koreman, Violeta Martínez-Paricio, Jardar Eggesbø Abrahamsen, Olaf Husby

Department of Language and Literature, NTNU, Trondheim, Norway
{jacques.koreman; violeta.martinez-paricio; jardar.abrahamsen; olaf.husby}@ntnu.no

ABSTRACT

This article describes an extension of the Computer-Assisted Listening and Speaking Tutor (CALST), an online pronunciation training platform. New exercises help L2-learners of Norwegian to overcome repairs caused by the violation of the phonotactic constraints of their native language. The article presents the contrastive analysis implemented for this purpose and discusses the repair strategies used by speakers when they learn a new language. It also discusses the limitations of the approach and the opportunities for learning from the errors which learners make.

Keywords: pronunciation training, phonotactics, consonant clusters, contrastive analysis, multilingual

1. INTRODUCTION

Based on the Swedish VILLE system [20], a free online platform for vocabulary and pronunciation training was developed for Norwegian dialects [2, 10, 12]. The platform is called the Computer-Assisted Listening and Speaking Tutor (CALST). Although exercises in CALST are presently only available for Norwegian, the platform is devised such that new content can easily be added for other languages, making existing exercise types immediately available for any new language. In that sense, CALST can be considered as a multilingual computer-assisted pronunciation training (CAPT) system.

The second novelty of CALST lies in the adaptive learning trajectory. Exercises are selected into the learning trajectory taking into consideration the specific needs of the learners, which vary dependent on their specific native language. It is well known that the acquisition of the sound system of a second language (L2) is a very complex process where different factors interact [1]. Yet, it has often been observed that the native language of the learner (L1) plays an important role in the process: sounds and sound contrasts absent in the learner's L1 generally pose a challenge. To help learners overcome their difficulties, pronunciation exercises for practising listening, speaking and writing skills have been developed in CALST based on a contrastive comparison of their L1 and the target L2 they wish to acquire. This contrastive analysis is carried out automatically in L1-L2*map*, an open-access, online system for language comparison which at present contains the phoneme inventories of over 500 languages [14, 11]. The phoneme inventories are an extension of the material available in UPSID [15, 16] and LAPSyD [17]. Using the output from the L1-L2*map* comparison, pronunciation exercises can be selected so as to make the learning process as effective as possible, paying particular attention to the training of the most difficult and unfamiliar sounds for each learner.

To improve the pronunciation and listening training in CALST, the system has recently been extended with exercises that deal with differences in the phonotactic constraints active in different languages. This paper discusses the methodology, challenges and results of our contrastive approach to pronunciation training.

The paper is organized as follows. In section 2, common repair strategies which are the result of phonotactic restrictions in the learner's L1 are discussed, as well as the exercises to overcome these repair strategies. Section 3 presents our methodology for predicting phonotactic challenges, describing three different solutions. In section 4, unresolved phonotactic problems are presented. The article finishes with conclusions and an outline of our research plan for the immediate future.

2. KNOWN CONSONANTS, NEW CHALLENGES

On the basis of a contrastive analysis in L1-L2*map*, CALST offers L2-learners exercises for speech sounds that are not part of the phoneme inventory of the learner's native language. However, to master the pronunciation of an L2 and avoid a foreign accent, it is not sufficient to learn the unfamiliar sounds of a language. The learner must also learn the particular combinatorial restrictions of each of these sounds in the L2, which often differ from those active in their L1. Even if a given speech sound exists in the learner's L1, it may be subject to different phonotactic restrictions than in the L2, and this may present a challenge to the learner [21, 1]. Several repair strategies may be used by learners to deal with violations of the L1 phonotactic constraints.

2.1. Common repair strategies

The repair strategies that learners use are not predictable from their L1. At the same time, the repair strategies that L2 learners apply are well known.

2.1.1. Syllabic restrictions

Even when a phoneme occurs in both the L1 and the L2, learners may find it difficult to pronounce when it occurs in a different syllable position from their L1. Well-known examples are the pronunciation of final voiced consonants in an L2 by speakers of an L1 that applies final devoicing, which blocks voiced obstruents from syllable-final positions even if they are allowed in the syllable onset. Another example is the mispronunciation of /r,l/ in a syllable nucleus by learners whose native language does not allow syllabic consonants (even if their L1 allows /r,l/ in the onset and/or coda). Likewise, the pronunciation of coda consonants can be a challenge for speakers whose L1 allows only few consonants in this position (Mandarin Chinese) or none at all (Tukang Besi). Regardless of their familiarity with the L2 consonants, learners of such languages need to practise them in unfamiliar positions in the L2.

2.1.2. Consonant clusters

Besides providing training on the distribution of individual consonants (previous section), the implementation of phonotactic constraints in our CAPT platform also deals with combinatorial restrictions on consonants in the L2. Some languages allow larger maximum syllable templates than others. While Tukang Besi for example only allows (C)V syllables [5], other languages allow complex consonant clusters in the onset and coda, as in Polish. Clearly, the pronunciation and perception of complex syllable structures may present a challenge to speakers of languages with only simple syllable structures.

Languages differ not only in the size of the syllable template they allow, but also in the particular content of such templates. For instance, a potential L1 and L2 might share the maximum syllable template CCV (i.e. the biggest syllable these languages allow contain a maximum of two consonants in the onset, and they do not permit any coda consonants). However, if the L1 only allows complex onsets consisting of a stop followed by a glide, but the L2 permits other types of complex onsets (e.g. stop+liquid, stop+fricative), the learner will have to practise the production and perception of such consonant clusters. Another important difference between languages is the degree to which they adhere to the Sonority Sequencing Principle (SSP) [18, 3]. This principle requires that syllables rise in sonority towards the nucleus and decrease in sonority from the nucleus. Additionally, languages may impose different intrasyllabic minimal sonority distances on their syllables. Although languages do not always adhere to the SSP, sonority restrictions active in the L1 often play an important role in the acquisition of an L2 sound system.

For instance, some languages may require a greater sonority distance between neighbouring consonants or an agreement of voicing between consecutive consonants within a syllable. In addition, languages sometimes do not allow specific clusters (e.g. */tl/ in the syllable onset in English and European Spanish), even if other sequences of consonants with similar sonority relations are allowed, e.g. /pl/ or /tr/. Learners will have to practise the production and perception of unfamiliar consonant clusters.

2.2. Overcoming repair

To deal with consonants clusters in the L2 which violate the phonotactic constraints of their L1, learners use repair strategies to accommodate unfamiliar structures and combinations to the L1 phonological system. One common strategy is simplification of syllable structures by deleting a consonant, for example the deletion of a coda consonant from a CC-cluster by speakers of an L1 that does not allow coda clusters. Although there is a general preference for coda consonants to be sonorous, for instance, this is not always the case. As a result, Vietnamese learners of English sometimes reduce clusters consisting of a liquid followed by a stop to a stop, although they generally produce a liquid instead [7]. Other strategies to adapt L2 structures to L1 constraints include the use of replacement of one sound with another and metathesis of consonants.

To train learners in CALST to overcome these repairs, we have developed ABX and minimal-pair exercises. These exercises parallel the use of such exercises for training simple sound contrast in CALST (please see the CALST website). In ABX exercises, a learner may hear [tɑlk] - [tɑk] (E. ‘soapstone’ and ‘thanks’, respectively) followed by the repetition of one of the two words. The learner then has to decide which of the two words the last word corresponds to. ABX exercises help learners to focus on the acoustic realization of the coda. In minimal pair exercises, the learner only hears one word and has to decide which of two words (e.g. <talk> or <takk>) it corresponds to. This exercise requires that the learner has internalized the correct pronunciation of the words. In addition to listening

exercises, pronunciation and spelling exercises are also provided, using the same materials.

Another well-known repair strategy for dealing with consonant clusters which violate L1 constraints is vowel epenthesis. By inserting a vowel before a complex onset (prothesis), a consonant cluster which is illegal in the learner's L1 can be broken up to two syllables which are permissible in the L1. An example of this is prosthesis of /ɛ/ by Spanish or Farsi learners of Norwegian before an sC- complex onset (sCV > [ɛ]s.CV). Other learners, e.g. speakers of Japanese, may use vowel epenthesis between the consonants in a cluster (anaptyxis) or after an illegal coda consonant to resolve similar problems. Since it is impossible to find sufficient minimal pairs to train learners to avoid epenthesis, training for this repair strategy in CALST is limited to self-evaluation exercises where learners can compare their own pronunciation with that of a tutor. Since it is well-known that learners may be insensitive to their mispronunciations (cf. [13]), our aim is to incorporate automatic speech recognition techniques to evaluate the pronunciation of consonant clusters in the future.

3. PREDICTING REPAIRS

Repairs are often applied when the phonotactic constraints of the learner's L1 are violated by L2 structures. In our contrastive approach, we attempt to predict when these occur. This is not in itself a difficult task: repairs often occur when there are stronger phonotactic constraints in L1 than in L2. Ideally, one needs a full list of all onsets and codas to predict possible repairs. Since these are often not available, we discuss different ways of dealing with this lack of data in this section.

3.1. Predicting phonotactic problems

We are not aware of any multilingual databases which contain detailed information listing all possible onsets and codas per language. Consequently, we have had to implement this ourselves in the L1-L2*map* database so as to enable comparison of the consonant clusters in L1-L2 pairs. The methodology that we adopted is the following.

First, for ten languages in the L1-L2*map* database, including Norwegian, full onset and coda cluster inventories have been created on the basis of the literature about these languages. The languages were selected on the basis of linguistic diversity and size of the immigrant group in Norway. Second, by comparing Norwegian consonant clusters with those of the other languages, a list of unfamiliar clusters can be created online for each of the languages. These lists are automatically linked to exercises in CALST. In that way, a learner whose native language is Mandarin Chinese, English or Spanish will be exposed to different exercises on Norwegian phonotactics, based on the different phonotactic constrains in the L1.

For the other languages in the database, no list of admissible clusters is available yet. Since L1-L2*map* is a wiki, it is possible for language experts to manually add a list of possible onsets and codas for the language(s) which they are interested in.

Ideally, we should like to use the maximum syllable templates available in the LAPSyD [17] and StressTyp2 databases [6], expanding them to cluster inventories in a principled manner. When information about specific languages is lacking, it would at first glance be tempting to use our knowledge about the universal tendencies discussed in Section 2.1.2 to create lists of syllable templates relevant for exercises on the basis of the phoneme inventories of the language. However, these tendencies are not real (hard) universals and have different priorities in different languages (which would correspond to different rankings of violable constraints in Optimality Theory), and they interact in different ways. At a detailed level, for instance, even two languages that permit complex onsets, may differ in the permitted sonority distance between the onset segments, and this variation furthermore interacts with other variables. As a result, a total expansion of syllable templates relevant for the exercises would instead turn into an unmanageable amount of consonant clusters and predicted repair strategies that in most cases would be irrelevant for the specific L1. Therefore a more pragmatic approach will be pursued instead, to be described in the following section.

3.2. Learning from learners

3.2.1. Observing phonotactic problems

CALST makes use of the learners' logged data in order to show a progress report on completed exercises. Both time and test results are logged. The result logging enables the system to learn what syllable structures are not problematical for learners with a given L1. For an initial set of learners, all exercises (consonant clusters and repair strategies) are offered as part of the learning path. After a minimum number of learners with a shared L1 have completed an exercise for a given onset or coda structure, and all of them have made no or only few mistakes, the onset or coda can be added to the list of unproblematic onset and coda clusters for that L1 (note that some onsets or codas may be unproblematic even though they are not part of the L1, for instance

because they are universally unmarked) and the particular exercise is taken out of the exercise list for new learners with the same L1, leaving only exercises which learners find difficult in the learning path. In this way, CALST learns from the learners, and the resulting knowledge is used to create a more efficient learning path through the exercises.

3.2.2. Observing repairs

In her descriptions of loanword phonology, which often parallels phonotactic adaptation in L2 acquisition, Smith argues that repairs may vary depending on whether the loans are auditory or orthography-based [19]. Also, it has often been attested that speakers of different L1s may use different epenthesis strategies (prosthesis, anaptyxis) for different consonant clusters (e.g. [22]). Since the selection of a repair strategy is unpredictable, all repair strategies must be implemented in CALST and offered to all learners, irrespective of their L1 (but see examples on cluster simplification in the first part of section 2.2).

The result logging in CALST not only allows us to evaluate which onsets and codas are problematic for learners with a given L1 (see Section 3.2.1), it also allows us to evaluate which repair strategy/ies they apply, since these will lead to errors in the exercises. CALST thus offers a solution to the apparent impossibility of predicting on the basis of their L1 which repair strategies learners use.

Importantly, the logged data can in future also be used for linguistic analysis. They can corroborate or falsify claims or predictions made by phonological theories of markedness or sonority, for instance. Since we hope to extend CALST to other languages than Norwegian, the platform can create new data using a strategic approach to L2 acquisition.

4. DISCUSSION: UNRESOLVED PROBLEMS

4.1. Limitations of a simple contrastive analysis

On the basis of a contrastive analysis, unfamiliar sounds in the L2 can be determined. But even if a sound is familiar, this does not mean that it will be pronounced correctly by all learners. We can draw a parallel to English loanwords in Hindi, where aspirated stops are replaced by unaspirated stops, while aspirated stops are used as substitutes for voiceless fricatives, which are not part of the Hindi language [9]. Regardless of whether they do so on the grounds of acoustic dissimilarities or to use native-language features to maintain phonological distinctions between all sound classes, the substitution (assuming it may also occur in L2 acquisition) will not be predicted by a simple contrastive analysis and no exercises will be offered in CALST to train correct pronunciation of a (familiar, or very similar) sound in L2. Clearly, a more sophisticated approach is needed to predict this type of substitution.

4.2. Consonant cluster dependencies

In order to limit the complexity of the exercises, our analysis focusses on consonants (clusters) in the syllable onset and coda. Dependencies between the nucleus and the coda do of course exist: For instance in Spanish one more consonant is allowed in the coda after a short vowel than after a diphthong [8].

CALST also does not consider the possible dependencies between the onset and the coda (see for example [4]). Since we are not familiar with research on the effects of these dependencies on L2 acquisition, we have for the time being ignored them and, in accordance with [3], assume that the demisyllable is the relevant unit in L2 learning.

5. CONCLUSIONS AND EXPANSIONS

This paper has examined the current expansions of CALST, an online platform which allows learners of an L2 (in this case, Norwegian) to practise their pronunciation and listening skills. We have seen that CALST's main novelty is its ability to take into account the properties of the L1 of each learner. Besides focussing on training the acquisition of new sounds, CALST now also provides exercises to practise the production and perception of consonant sequences and L2-particular phonotactics.

To avoid a strong foreign accent and ensure an effective communication, not only segmental and distributional properties must be learned, but speakers must also acquire the new prosodic system of the L2: the position of stress and its acoustic-phonetic cues, the tonal patterns of languages with lexical tone, and intonational properties of languages. In the near future, we hope to extend L1-L2*map* and CALST with information about the position of stress, making use of the StressTyp2 database [5], which contains information about stress in a large number of languages. Likewise, we would like to incorporate exercises to train learners of Norwegian to perceive and produce the lexical tones and pitch-accent patterns of the language.

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

6. ACKNOWLEDGEMENTS

We thank Egil Albertsen, Keivan Hedayatfar and Øyvind Bech for their contribution. The work reported on in this article was supported by Vox, the Norwegian Agency for Lifelong Learning.

7. REFERENCES

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

[2] CALST, The Computer-Assisted Listening and Speaking Tutor. http://calst.no.

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

[4] Davis, S., Baertsch, K. 2011. On the Relationship between Codas and Onset Clusters. In: Cairns, C., Raimy, E. (eds), *Handbook of the Syllable.* Leiden: Brill, 71-97.

[5] Donohue, M. 1999. *A Grammar of Tukang Besi.* Berlin: De Gruyter Mouton.

[6] Goedemans, R., Heinz, J, van der Hulst, H. 2014. StressTyp2. http://st2.ullet.net.

[7] Hansen, J.G. 2004. Developmental sequences in the acquisition of English L2 syllable codas: A preliminary study. *Studies in Second Language Acquisition* 26(1), 85-124.

[8] Harris, J.W. 1983. *Syllable Structure and Stress in Spanish. A Nonlinear Analysis* (series: Linguistic Inquiry Monographs 8). Cambridge, MA: MIT Press, 1-158.

[9] Hock, H. 1988. Principles of Historical Linguistics. Berlin: Mouton de Gruyter.

[10] Husby, O., Øvregaard, Å., Wik, P., Bech, Ø., Albertsen, E., Nefzaoui, S., Skarpnes, E. & Koreman, J., 2011. Dealing with L1 background and L2 dialects in Norwegian CAPT. *Proc. of the workshop on Speech and Language Technology in Education (SLaTE2011),* Venice.

[11] 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.

[12] Koreman, J., Wik, P., Husby, O., Albertsen, E. 2013. Universal contrastive analysis as a learning principle in CAPT, *Proc. of the workshop on Speech and Language Technology in Education (SLaTE 2013),* 172-177.

[13] Kuhl, P.K., Tsao, F-M., Liu, H-M. 2003. Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Science* 100 (15), 9096–9101.

[14] L1-L2*map,* A Multi-Lingual Tool for Contrastive Analysis, http://calst.no/L1-L2map.

[15] Maddieson, I., 1980. UPSID: UCLA phonological segment inventory database. Phonetics Laboratory, Department of Linguistics.

[16] Maddieson, I. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.

[17] 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.

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

[19] Smith, J.L. 2006. Loan phonology is not all perception: Evidence from Japanese loan doublets. In Vance, T.J., Jones, K. (eds), *Japanese/Korean Linguistics* 14. Stanford: CSLI.

[20] Wik, P. 2011. *The Virtual Language Teacher*. Ph.D. thesis, KTH School of Computer Science and Communication.

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

[22] Zuraw, K. 2007. The role of phonetic knowledge in phonological patterning: corpus and survey evidence from Tagalog infixation. *Language* 83(2), 277-316.