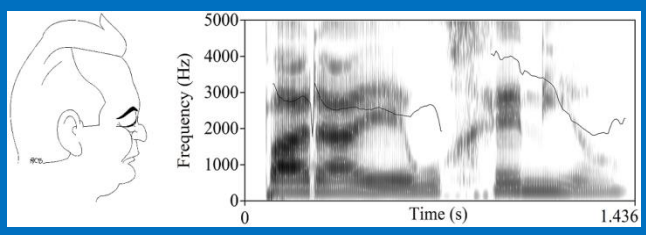
Perception and production of L2 speech sounds: L1 phonetic/phonological categories do not tell it all

Lucrecia Rallo Fabra



Fernández Planas, A. Ma. (ed.) (2016): *53 reflexiones sobre aspectos de la fonética y otros temas de lingüística*, Barcelona, págs. 417-425.
ISBN: 978-84-608-9830-6.

Perception and production of L2 speech sounds: L1 phonetic/phonological categories do not tell it all

Lucrecia Rallo Fabra
Universitat de les Illes Balears
lucrecia.rallo@uib.es

Thank you Eugenio for your mentorship as a tutor and PhD supervisor. Thanks for giving me the chance to work at the Phonetics Lab as an assistant researcher. Thanks for opening my eyes to the universe of speech perception.

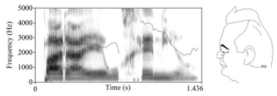
1. EARLY WORK IN SPEECH PERCEPTION AT THE PHONETICS LAB

During my first years working as an assistant researcher at the UB Phonetics Lab in the early nineties, the Spanish Ministry of Education and Science funded a research grant on acoustic invariance in Spanish. This was my first contact with speech perception and instrumental phonetics. At that time, speech visualisation and editing was complex and it was mostly done with the *DSPSonograph 5500*, an oven-size unit with many lights and buttons similar to the control panel of a spaceship coming from Mars. Running a perceptual experiment also required complicated logistics, such as, recruiting a large-enough group of undergraduate students and squeezing them into the cell-size sound-proof non-air-conditioned lab, where they would listen to bursts and noises and portions of speech sounds and subsequently make decisions about the phonetic identity of the sounds. Despite all these limitations, the research team lead by Prof. Martínez Celdrán managed to obtain interesting results, which were published in volume V of *Estudios de Fonética Experimental.*

This work was pioneering in Spain; to our knowledge, no other labs or university departments were carrying out this kind of research in the field of speech perception. The rationale that motivated the project was whether findings by Blumstein and Stevens (1979; 1980) about the existence of invariant acoustic cues to contrasting consonants could also be extended to Spanish. In a series of perceptual experiments with synthetic stimuli, Rallo and Fernández Planas (1995) investigated the role of various acoustic cues, namely, release burst, vowel stimuli and vowel transitions on the identification of Spanish voiceless stops /p/, /t/, /k/. Results showed that consonant identity was easily perceived from synthetic stimuli as short as 10 or 20 milliseconds, suggesting that stimulus length was not determinant for correct consonant identification. In most cases, vowel transitions had a facilitating role to correctly identify the three Spanish voiceless stops. Interestingly, if the release burst was removed, listeners were still able to identify the consonant with relative ease, relying on vowel transitions alone.

In line with Strange (1995), we could attest the language-specific nature of phonetic category boundaries, so our next endeavour was to further investigate the acoustic characterization of speech sounds in cross-linguistic environments. In a cross-linguistic

 ISBN 978-84-608-9830-6



study about the effects of speaking rate on the production of English and Catalan /t/ in the word *tapo* (Rallo Fabra, 1997), acoustic measurements of overall word duration, VOT and closure interval revealed that VOT was positively correlated with word duration in English but not in Catalan. Conversely, closure intervals were positively correlated in Catalan but not in English. These findings showed evidence of a distinct cross-linguistic difference in the way speech rate was acoustically implemented in both languages. While English speakers shortened VOT duration as speech rate increased, Catalan speakers would keep VOT constant across rates and shorten closure intervals as speech rate increased.

In a follow-up study (Rallo Fabra, 1999), we addressed the question of whether L1-Catalan experienced learners of English would transfer the language-specific patterns of Catalan and thus produce English /t/ with short VOT durations, or whether their VOTs would reach a compromise between Catalan learners. The results revealed that the VOT values obtained by the Catalan learners were intermediate between both languages. As for speech rate, we found high intra-learner variability in terms of how speech rate affected the production of English /t/. Some learners followed the L1 pattern found in Catalan and increased closure duration as speech rate increased, others modelled the English pattern, increasing the duration of VOT. However, most learners increased both acoustic cues, VOT and closure interval, indicating a compromise pattern between both languages. A question that remained unanswered was whether learners had modified their previously established L1 phonetic/phonological categories to accommodate Catalan and English /t/ or whether they established a new separate category for English /t/.

2. THE ROLE OF PHONETIC/PHONOLOGICAL CATEGORIES IN L2 SPEECH PERCEPTION AND PRODUCTION

The findings of the studies just reported can be accounted for by the predictions and hypotheses of two theoretical models that have motivated a great deal of empirical research on L2 perception and production during the last two decades, namely, the Speech Learning Model, SLM for short (Flege, 1995) and the Perceptual Assimilation Model, (PAM) and its extension, PAM-L2 (Best and Tyler, 2007). Both models are based on the assumption that non-native speakers of a given language perceptually assimilate the L2 categories to their L1. The SLM assumes that this assimilation operates at the phonetic level and distinguishes between similar, new and old sounds, depending on the degree of similarity between the L2 categories and their closest L1 counterparts. Success in perceiving and producing the new L2 categories will be guaranteed if learners establish new phonetic categories for the new L2 sounds.

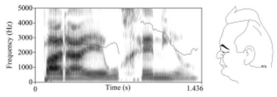
The PAM-L2 makes predictions of L2 discriminability based on four different patterns of perceptual assimilation of L2 categories to L1 categories. Thus, the L2 contrast can assimilate to one L1 category (single-category assimilation), two L1 categories (two-category assimilation), or it cannot assimilate to any L1 category (uncategorized). This model assumes that the L1 and L2 categories interact at both the phonetic and phonological levels. This involves one single interlanguage phonological category which serves the two languages and is realized as two separate L1 and L2 phonetic categories.

Having the SLM and the PAM-L2 as frameworks, Rallo Fabra and Romero (2012) investigated the perception and production of American English vowels by three groups of advanced Catalan learners of English varying in language proficiency. Learning to perceive and produce the English vowel system is a challenge for Catalan speakers, because the vowel system of standard American English has thirteen vowels (Peterson and Barney, 1952), as opposed to Central Catalan (Recasens and Espinosa, 2006), which has only eight vowels. In the light of the asymetries between the two languages, we administered two discrimination tasks, one testing Catalan-English vowel contrasts /a-ʌ/, /a–ɑ/, /ɛ–ɛ/, /a–ɛ/, /a–æ/, /i–i/, /i–ɪ/, /u–u/, and another task testing English vowel contrasts /ɑ–ʌ/, /ɛ–æ/, /i–ɪ/, /u–ʊ/. Listeners' sensitivity to the different vowel contrasts was measured in terms of an A', a proportion of hits and false alarms derived from each contrast (Snodgrass and Levy-Berger, 1985). Overall, listeners distinguished the C–E /i–i/ contrast relatively well, and they could also distinguish between the Catalan and English vowels /i–i/, /u–u/, and /a–ɑ/. However, the Catalan–English /a–e/, /a–æ/ and /e–e/ vowel contrasts proved extremely difficult to discriminate for the non-native listeners. As for the English–English pairs, the learners could discriminate the speech sounds in the /i– ɪ/ and /u–ʊ/ pairs, suggesting that, as predicted by the SLM, the Catalan learners had probably established new phonetic categories for the English vowels /ɪ/ and /ʊ/.

The low perceptual sensitivity shown by the Catalan listeners in relation to the contrasts involving the English low vowels /ʌ/, /ɑ/, /æ/ could be accounted for by the predictions of the PAM framework, which states that when two L2 sounds are perceptually assimilated to one single L1 category, discrimination is difficult. In our case, Catalan listeners were faced with three English low vowels /ʌ/, /ɑ/, /æ/, which were assimilated to one single L1 category /a/. We hypothesized that learners could not establish new phonetic categories for these three English vowels, which probably merged in their perceptual vowel space with its Catalan counterpart /a/.

However, phonetic and phonological categories do not tell it all. As Best and Tyler (2007) point out, as learners make progress in the learning ladder, they pay less attention to low-level phonological information and switch their attention to higher-order linguistic levels, such as semantics or syntax. From the psycholinguistics perspective, Trofimovich (2005, 2008) further attests that attention to meaning interferes with learners' processing of the phonological forms of words. L2 learners get distracted by differences in speakers' voices and are unable to focus on the important phonetic properties that signal the identity of the spoken word. This raises the question of whether exposure to a wide range of speakers and dialects has a facilitative role in establishing these new perceptual categories.

Recent research in L1 infant speech perception has demonstrated that exposure to a wide range of phonetic variability in terms of L1 accents and speakers is paramount to grasp phonological constancy i. e. the ability to recognize familiar words pronounced by speakers from different dialects. In this line, Best and cols. (Best et al. 2009; Mulak et al, 2013) investigated the role of accent variability in toddler's sensitivity to familiar vs. unfamiliar L1 accents. Through an eye-tracking task 19-month-olds looked longer at images spoken in an unfamiliar accent indicating that phonological constancy emerged by 19 months and that its development enhanced vocabulary growth and later reading acquisition.



In L2 acquisition, the effect of exposure to diferent L2 accents has received little attention. Escudero and Chladkova (2010) found that accent influenced the way Spanish listeners assimilated English vowel categories into Spanish vowel categories. The assimilation patterns changed depending on whether the vowel sound they heard was from BrEng or from AmEng. These findings are in line with earlier work examining the perceptual assimilation of English vowels to Catalan vowels by Catalan listeners with different L2 experience (Cebrian, 2006; Rallo, 2005), showing that many factors interact in the perceptual mapping of L2 sounds onto L1 categories, among them, listeners' experience with the L2, the consonant context of the vowel stimuli used in the experiments, and the English dialects spoken by the speakers. According to Strange (1998), when perceiving vowel stimuli, native and non-native listeners are not equally sensitive to the phonetic variation found in different instances of the same vowel category. It seems that he non-native listener pays more attention to dialect and individual within-category differences of the tokens, which are largely ignored by the native listener. Nevertheless, at present, it is uncertain whether exposure to cross-accent variability would have a facillitating effect for the establishment of L2 phonological categories and, if so, how broad that variability should be for optimal L2 speech development.

2.1. Factors influencing L2 speech learning

A substantial number of studies have provided supporting evidence that phonetic/phonological categories alone cannot fully predict the ease/difficulty of acquisition of L2 sounds. For instance, Amengual (2012) examined the production of VOT in four groups of Spanish–English bilinguals: Spanish heritage speakers, English heritage speakers, advanced learners of English and advanced learners of Spanish. He found that there was a strong effect of cognate status on the duration of VOT in English. Specifically, that non-native English speakers produced shorter and, thus, more Spanis-like VOTs in cognates (words with close phonologies and same meaning in both languages, i. e. *teléfono*–'telephone'), than in non–cognates (words with different phonologies and meanings in both languages), suggesting that these speakers probably had the same mental representations for the English and Spanish cognates. Along the same line, Mora and Nadeu (2012) reported similar cognate-status effects in vowel production by Catalan-Spanish bilinguals. Language dominance influenced production of the Catalan vowel /ɛ/ by Spanish-dominant bilinguals, who produced instances of Catalan /ɛ/ which were closer to Spanish /e/. In contrast, production of /ɛ/ in non-cognates was easier because bilinguals could not transfer the pronunciation of a word that does not exist in Spanish.

A recent study examining pronunciation of English words by Catalan/Spanish learners in classroom settings (Rallo Fabra and Jacob, 2015), further attests that sucess/failure in learning the sounds of a non-native language cannot only be predicted by a missmatch between the L1 and L2 phonetic/phonological categories. In this study we found that orthography played an important role in the production of English vowels by Spanish Catalan intermediate learners, who were recorded while reading aloud a story. It is well known that the three languages differ in terms of orthographic transparency, whereas Spanish and Catalan can be considered «transparent languages» in the sense that there is a close connection between the sounds and their spelling, English is considered an «opaque language» in the sense that the same sound can be represented with different spellings

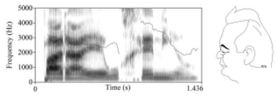
(i.e. the English vowel /ɪ/ can be represented orthographically with the spellings *i, y, e, a, o, ui, u* as in: *big, system, pretty, image, women, building, business*. Overall, the Spanish/Catalan learners produced more native-like instances of the English vowels when the spelling of the target words was almost identical to the sound, as in the word big. However, when there was a mismatch between the sound and the spelling of the target word, the number of pronunciation errors increased. These findings supported previous studies in which reading instruction and phoneme-to-grapheme conversion rules were found to negatively influence children's perceptual abilities. As children's reading skills improved, their perceptual sensitiveness to non-native speech sounds decreased gradually (Burnham et. al. 2002). In the case of adults, Escudero (2015) has also shown that orthography may not facilitate L2 speech learning and suggests that avoiding orthographic input could be a good strategy to follow in formal language learning environments.

2. 2. Questioning the role of phonetic/phonological categories: exemplar models

According to Port (2007), the complex process of language places a heavy load on human memory, which entails that language must be encoded in some highly efficient way. It follows that memory stores much information in the form of specific events rather than in an abstract phonological form. This is why he questions the widely accepted assumption that phonetic features are what speakers use to build the words stored in the lexicon. Arguably, they use a «rich and detailed description of words that combines linguistic and nonlinguistic properties of recently heard speech signals». The non-linguistic properties would include speaker-dependent detailed information about voice quality and speech among others.

The term «exemplar memory» is thus used to refer to a memory that stores concrete examples or «clusters of neighbors of various kinds in the speech auditory space». From this perspective, the phonology of a language would be understood as a set of overlapping categories of similar words in the memory of speakers. Port further argues that words may also be discrete, just like phonemes. Somehow apparently, the database of tokens of individual speech fragments (such as words) is able to influence a speaker's choice of pronunciation decisions, since speakers (especially younger ones) modify their pronunciations to be more similar to what they hear others say, as it had been shown in earlier work (Goldinger, 1998; Pierrehumbert, 2001).

Recent work by Rallo Fabra and cols. has shown that the difficulties shown by bilinguals and L2 learners could be accounted for by the exemplar model just reported. For instance, in Rallo Fabra (2015), the errors shown by late Spanish–English bilinguals to reduce English unstressed vowels seemed to indicate that they did not abstract the English stress patterns across the lexicon, instead they probably learned these patterns on a word-by-word basis. In the case of L1–Spanish/L1–Catalan L2 learners, Cortés and Rallo Fabra (2016) suggested that learners' insertion of epenthetic *e* in words such as *speak, stop* and *skate* was influenced by lexical frequency, that is, less frequent words were prone to more epenthetic vowel insertions than more frequent words. Similarly, Rallo Fabra and Jacob (2016) found that the pronunciation difficulties faced by young L1–Spanish/L1–Catalan learners of English depended on both, the cognate status of the target words with respect



to Spanish or Catalan and the orthographic transparency of these words. A clear example of this is a learners' elicitation of English 'politician' as *policía*. We speculate that maybe this learner's lexicon did not include exemplars of the target English word and that his pronunciation decision when eliciting *policía* was misguided by the scaffolding role of orthography.

3. REFERENCES

AMENGUAL, M. (2012): «Interlingual influence in bilingual speech: Cognate status effects in a continuum of bilingualism», *Bilingualism: Language* and *Cognition*, 15, pp. 517–530.

BEST, C. T. and M. D. TYLER (2007): «Nonnative and second-language speech perception», in O. –S. Bohn and M. Munro (eds.): *Language experience in second language speech learning: In honor of James Emil Flege*, Amsterdam, John Benjamins, pp. 13–34.

BEST, C. T.; M. D. TYLER. T. N. GOODING. C. B. ORLANDO and C. A. QUANN (2009): «Development of phonological constancy: Toddlers' perception of native- and Jamai- can-accented words», *Psychological Science*, 20, pp. 539–542.

BLUMSTEIN, S. E. and K. N. STEVENS (1979): «Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants», *Journal of the Acoustical Society of America*, 66, pp. 1001–1017.

BLUMSTEIN, S. E. and K. N. STEVENS (1980): «Perceptual invariance and onset spectra for stop consonants in different vowel environments», *Journal of the Acoustical Society of America*, 67, 2, pp. 648–662.

BURNHAM, D. K; M. TYLER and S. HORLYCK (2002): «Periods of speech perception development and their vestiges in adulthood», in A. Rohde (ed.): *An integrated view of language development: Papers in honor of Henning Wode*, Trier, Wissenschaftlicher Verlag, pp. 281–300.

CEBRIAN, J. (2006): «Experience and the use of non-native duration in L2 vowel categorization», *Journal of Phonetics*, 34, pp. 372–387.

CORTÉS, S. and L. RALLO FABRA (2016): «Production of English consonant clusters in onset position by Spanish-Catalan speakers», poster presented at the *8th International Conference on Second-Language Speech, New Sounds*, Aarhus.

ESCUDERO, P. (2015): «Orthography plays a limited role when learning the phonological forms of new words: The case of Spanish and English learners of novel Dutch words», *Applied Psycholinguistics*, 36, 1, pp. 7–22.

FLEGE, J. E. (1995): «Second-language speech learning: Theory, findings, and problems», in W. Strange (ed.): *Speech perception and linguistic experience: Issues in cross-language research*, Baltimore, York Press, pp. 233–277.

GOLDINGER, S. D. (1998): «Echoes of echoes? An episodic theory of lexical access». *Psychological Review*, 105, pp. 251–279.

MORA, J. C. and M. NADEU (2012): «L2 effects on the perception and production of a native vowel contrast in early bilinguals», *International Journal of Bilingualism*, 16, 4, pp. 484–500.

MULAK, K. E; C. T. BEST, M. D. TYLER, C. KITAMURA and J. R. IRWIN (2013): «Development of phonological constancy: 19-month-olds, but not 15-month-olds, identify words in a non-native regional accent», *Child Development*, 84, 6, pp. 2064–2078.

PETERSON, G. E. and H. L. BARNEY (1952): «Control methods used in a study of the vowels», *Journal of the Acoustical Society of America*, 24, pp. 175–184.

PIERREHUMBERT, J. B. (2001): «Exemplar dynamics: word frequency, lenition and contrast», in J. L. Bybee and P. L. Hooper (eds.): *Frequency and the emergence of linguistic structure*, Amsterdam, John Benjamins, pp 137–158.

PORT (2007): «How are words stored in memory? Beyond phones and phonemes», *New Ideas in Psychology*, 25, 2, pp. 145–172.

RALLO, L. and A. MA. FERNÁNDEZ PLANAS (1995): «La invariación acústica en las oclusivas sordas del castellano. Estudio perceptivo», *Estudios de Fonética Experimental*, VII, pp. 45–74.

RALLO FABRA, L. (1997): «Speaking rate effects in Catalan and English. A cross-language study», *Estudios de* Fonética *Experimental*, IX, pp. 243–272.

RALLO FABRA, (1999): «Speaking rate effects in voiceless stops produced by Catalan speakers of English», in J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville and A. C. Bailey (eds.): *Proceedings of the 14th international congress of phonetic sciences*, Berkeley, Regents of the University of California, pp. 1417–1420.

RALLO FABRA (2005): «Predicting ease of acquisition of L2 speech sounds. A perceived dissimilarity test», *Vigo International Journal of Applied Linguistics*, 2, pp. 75–92.

RALLO FABRA (2015): «Can nonnative speakers reduce English vowels in a native-like fashion? Evidence from L1-Spanish L2-English bilinguals», *Phonetica*, 72, 3–4, pp. 162–181.

RALLO FABRA, L. and K. JACOB (2015): «Does CLIL enhance oral skills? Fluency and pronunciation errors by Spanish-Catalan learners of English», in M. Juan–Garau and J. Salazar–Noguera (eds.): *Content-based language learning in multilingual educational environments*, Cham, Springer, pp. 163–178.

RALLO FABRA, L. and K. JACOB (2016): «When 'politician' becomes 'policía' and 'border' becomes 'brother': effects of orthography and cognate status on Spanish-Catalan EFL learners' pronunciation», poster presented at the *8th International Conference on Second-Language Speech. New Sounds 2016*. Aarhus.

RALLO FABRA, L. and J. ROMERO (2012): «Catalan learners' perception and production of English vowels», *Journal of Phonetics*, 40, 3, pp. 491–508.

RECASENS, D. and A. ESPINOSA (2006): «Dispersion and variability of Catalan vowels», *Speech Communication*, 48, PP. 645–666.

SNODGRASS, J; G. LEVY-BERGER and M. HAYDON (1985): *Human experimental psychology*, New York, Oxford University Press.

STRANGE, W. (1995): «Phonetics of second-language acquisition: Past, present, future», in K. Elenius and P. Branderud (eds.): *Proceedings of the XIII international congress of phonetic sciences*, Stockholm, Arne Stombergs, pp. 76–83.

STRANGE, W; R. AKAHANE-YAMADA; R. KUBO; S. A. TRENT; K. NISHI and J. J. JENKINS (1988): «Perceptual assimilation of American English vowels by Japanese listeners», *Journal of Phonetics*, 26, 4, pp. 311-344.

TROFIMOVICH, P. and K. MCDONOUGH (2008): *Applying priming methods to L2 learning, teaching and research. Insights from psycholinguistics*, Amsterdam, John Benjamins.