UCLA

Voices

Title

Why Teaching Pronunciation to Spanish L2 Learners Matters

Permalink

https://escholarship.org/uc/item/51g604gx

Journal

Voices, 2(1)

Author

Meléndez-Ballesteros, Nancy

Publication Date

2014

Copyright Information

Copyright 2014 by the author(s). This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at https://creativecommons.org/licenses/by-nc-nd/3.0/

Peer reviewed

Why Teaching Pronunciation to Spanish L2 Learners Matters

Nancy Meléndez-Ballesteros University of California, Los Angeles Centro de Estudios del Español de Estados Unidos (CEEEUS)

Abstract

In an increasingly globalized world, second language learners need to learn how to communicate effectively and confidently. In this context, pronunciation is crucial. In this paper, I show that placing emphasis on form in a classroom environment helps with the perception and the production of a more native-like L2 (in this study, Spanish).*

Keywords: accuracy, audio-visual imitation tasks, focused-attention, fluency, native-accent pronunciation, perception-production

I. Introduction and Background. Learning a second language entails the accumulation of an unfamiliar linguistic system that may or may not be similar to that of the second language (henceforth, L2) learner. It is assumed that learners carry within the language storage area(s) specific linguistic knowledge of their first language that could interfere with the proper or full attainment of a second language.

Over the course of the past two-three decades, much of the research on L2 acquisition and second language instruction has mostly focused on the teaching of general aspects of an L2 (i.e., grammar, vocabulary, orthography, etc.) without paying much attention to Pronunciation. This has inadvertently contributed to the non-or improper-attainment of the L2 phonological system; thus, introducing in the Spanish-speaking world or community, speakers with moderate to heavy accented speech.

75

^{*} I would like to thank Dr. Claudia Parodi and the members of the Centro de Estudios de Español de Estados Unidos (CEEEUS) for their helpful comments and suggestions. In addition, a special thanks to Belén Villarreal, Bryan Kirschen, Lara D. Rann, and Dr. Montserrat Reguant for taking their time to review my paper.

This pilot study provides an overview of a more elaborate inquiry that is in progress. I will begin with a brief description of the differences between first and second language learners, followed by a background discussion on why perception should be linked to production, which in addition can help describe the challenges the L2 learner faces. Moreover, general reasons as to why pronunciation should be emphasized in L2 instruction will be given, with an idea on the type of listening-speaking training exercise that can be implemented in order to aid the L2 learner achieve or approximate a more native-like production of the L2 in question (in this study, Spanish).

2. The LI & L2 Learners: A Brief description. In order to understand what a second language learner faces when learning a second phonology, it is necessary to know how the first language is acquired. The L1 learner fully acquires his mother tongue during his early childhood years by the input s/he constantly receives from his or her surroundings. S/he starts the process of learning the mother's tongue with a clean slate. There is no accumulated or stored information in his memory; unlike our L2 learner, he/she cannot make any judgments on the grammar being heard, analyze it, etcetera. The child has no basis for comparison. The L1 learner relies on universal processes just like other children or infants in the world. This particular learner begins with an attempt to imitate the sounds he/she hears from his environment (first the mother, father, siblings, other family members, family friends, etcetera) and eventually starts producing words accurately. The imitation of the new sounds being heard might take place immediately or at a later time (Meltzoff, 1999).

Ll Learner

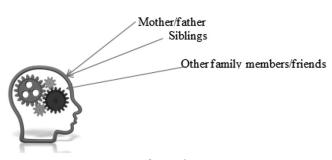


Figure 1. L1 learner's input

On the other hand, the L2 learner comes into a classroom setting with a language-specific system; a phonological system, to be exact, that contains very specific acoustic/phonological information that s/he will have to learn to ignore or set-aside in order to receive more openly the new one being heard. He or she needs to learn how to distinguish those phonological features that belong to the L1 and those that belong to the L2, which for most students learning a foreign language can be very confusing and frustrating. In other words, the L2 learner is dealing with two competing phonological systems: the one being introduced or heard in the L2 environment (the classroom), and the one that already exists within his/her phonological repertoire. The L2 learner needs to decide which one will be used when producing the L2. This should be done with a conscious effort as many L2 researchers have stated or implied, but as evidenced within many classroom settings, it is not.

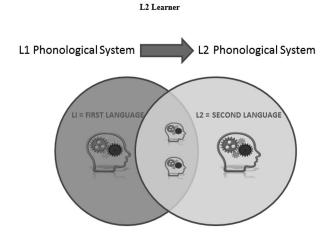


Figure 2. L2 learner's input

3. Background: Perception and Production. Deciding which phonological system the learner will eventually use, will depend on two things: the first, on the individual's perceptual acuity; the second, on the conscious effort he/she puts on learning the new phonological element(s), which usually is very low. I will refer to the learner's low learning effort with the Least Effort Approach, L.E.A. This approach follows a broad principle that has been used in many different fields, the Principle of Least Effort, which postulates that individuals when confronting a new task

will obviously choose the path with the least effort. In other words, he or she will take the easy way out. This idea can be easily carried over to the learning process of an L2 system, in this particular study, the phonology.

Students, when receiving an L2-speech signal, will generally choose the one that is closer or easiest to produce from their repertoire without focusing on the specific features. Thus, when producing a sound, the target will not be that of the L2 but rather an L1 sound that learners perceive as the most similar to the L2 phone that they hear. This of course brings us to the first issue, Perceptual acuity. When there is focused-ATTENTION, perceptual acuity helps the listener to discriminate the sounds from the incoming speech signal so that they can be easily understood and reproduced at a later time. For this reason, the idea of Perception has brought many researchers to concur that if an individual is unable to perceive or discriminate an L2 sound, s/he will not be able to produce it accurately² (Trubetzkoy, N.S., 1939; Barry, W., 1989; Flege, J. E., 1991, 1995). In addition, some studies have shown that the ability or inability to perceive these foreign sounds will depend on how similar or dissimilar the L2 phone is to learner's L1. The closer (similar) the L2 sound is to the one in the L1, the more difficult it will be to acquire, as it is suggested by a cognitive mechanism Equivalence classification that is part of Flege's (1992, 1995) Speech Learning Model (SLM). Others, on the other hand, have stated that it is not so much how DIFFICULT it is to acquire the L2 phones, but the RATE of acquisition. The dissimilar phones are acquired faster than similar ones as specified in the Similarity Differential Rate Hypothesis (SDRH) by Major & Kim (1996). Some have researched a little further and concluded that what some of the aforementioned models fail to do is demonstrate the WHEN and HOW this process of categorization of the similar and dissimilar phones will take place (Brown, 2000). This is precisely what L.E.A. and the idea of FOCUSED-ATTENTION will aim to explain in detail. Regardless of the approach taken, it is clear that the overall goal is to find one that forges a cogent link between PER-CEPTION and PRODUCTION.3

However, one must bear in mind that focusing on just perceptual hearing cannot possibly lead an L2 learner to become a proficient speaker of a second language. It is important to bring another essential sensory system that will be discussed briefly in a later section, namely, the visual. It is, therefore, the ultimate goal of this research to help students discriminate the sounds from the incoming speech signal by receiving

AUDIO-VISUAL input and IMITATING it; thus, being able to produce more native-like L2 phones. I propose that it is through the imitation tasks that the ultimate attainment of a more-native like production can be accomplished. These audio-visual tasks will help the L2 learners become more aware of the acoustic information or cues found within the articulatory gestures that can aid them to produce similarly to the speech model, whoever or whatever it may be. But, before delving into the details, it is necessary to explain the importance of including pronunciation (production) in an L2 curriculum.

4. Pronunciation: Why should it be brought back to L2 teaching? Within the last few decades, teaching methods have been implemented in the classrooms that have not rendered favorable results on the accurate or full attainment of the targeted phonological system. Many of these methods, whose goal is to teach general communication skills, do not take into account the diverse learning styles that L2 learners bring into a classroom. These have been considered to be too dogmatic and cannot be used with different learning settings (e.g., Brown, 2001; Kumaravadivelu, 2003, 2005).

Although a communicative based approach teaches a student to express his or her ideas with some type of fluency, to be somewhat understood, it does not teach him or her to focus on accuracy when producing the L2 sounds. As a result of this, the L2 learner leaves the classroom with a moderate to heavy accent. For most it ends there, but for those that would like to continue with foreign studies, use it in other fields (i.e., medical, business, translation, etc.), or be a part of the L2 community, will find that there will be communication difficulties due to intelligibility issues (i.e., Munro & Derwing, 1995).

It is a well-known fact, within the teaching field, that in order to be proficient in any language, the language learner should be both fluent and accurate in any aspect of the language (syntax, phonology, grammar, etc.) In this paper, I focus on the accuracy of pronunciation. If a student is said to be fluent, but not accurate, it means that s/he is able to manage the general aspects of a language (i.e. grammar, syntax, idiomatic expressions, etc.), but can have accent or accuracy issues that can obstruct the communication process. On the other hand, if the student is accurate, but not fluent, it means that s/he is able to produce without a noticeable accent, but there might be errors in the syntax, grammar, and so on. In either case, something must be done because, as was stated earlier, there

are many students within foreign language classrooms that are highly interested in learning how to speak the L2 clearly, with no noticeable accent, so that they can use it at a higher educational or professional level (i.e. business, medical field, etc.). Putting more emphasis on pronunciation can help this particular L2 learner decrease his/her foreign accent, which will eventually help minimize future miscommunications with his/her students, patients, business associates, and so on.

In this pilot study, in order to reduce and assess accentedness of speech, focus is placed on some phonemes that many (i.e. language instructors, students, etc.) consider to be the most difficult to attain, namely, the Spanish rhotics the tap /r/ [r] & the trill /rr/ [r], the Spanish voiced & voiceless plosives /b, d, g/ & /p/, /t/, /k, respectively. Acoustic measurements are typically used to determine whether an individual has produced the aforementioned sounds accurately or more native-like. For the latter, the Voice Onset Times (VOTs) (Lisker & Abramson, 1964) is measured; that is, from the release of the stop to the beginning of the following voiced segment, when the vocal cords begin to vibrate. This elapsed time, the time between the consonant release and the vibration of the vocal cords, is measured in milliseconds (ms) as depicted in FIGURE 3. As for the rhotics, the measurement is taken on the number of times the tongue has contact (taps) with the alveolar ridge as depicted in FIGURE 4, and the duration (the length of time it takes to produce the sounds). Preliminary results of the rhotics are given in this paper.

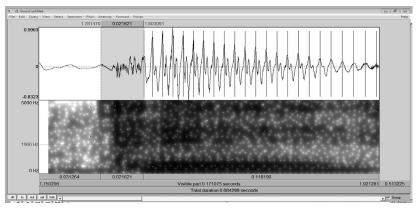


FIGURE 3. Non-native speaker's production of 'decir' (Plosive /d/). Only the first segment, /de/, is shown. The highlighted area is the segment that shows the release of the stop up to the beginning of the vowel.

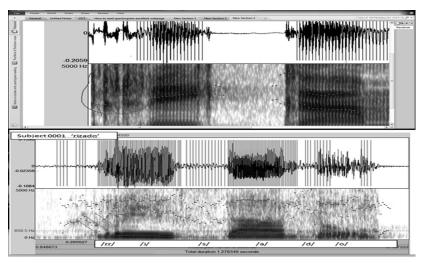


FIGURE 4. Typical numbers of taps for Spanish speakers, between 2 and 5. Top spectrogram depicts the subject's production of 'recámara' and bottom one shows the pronunciation of 'rizado'. The taps are indicated with green arrows for the top production, and with blue, for the bottom one.

5. Methodology

- 5.1 Participants. Seventy-four students, who were enrolled in an Elementary Spanish 2 course in a Southern California college during the spring 2012 semester, were given imitation tasks as part of their oral assignments. Thirty-seven of the students were enrolled online and thirty-seven in two different sessions within a traditional (in-class) setting. The majority of the participants were women, their ages ranged from 18 to 23 years. At the time of the implementation of the tasks and data collection, they had been taking Spanish for two semesters. There were non-native speakers of Spanish, of diverse ethnic backgrounds, and native (heritage) speakers of Spanish, the majority of Mexican descent.
- 5.2 Tasks. As stipulated in the syllabi, students are required to complete an extra hour of online homework per week. This hour is completed by using online voice tools with a program that is used by many instructors, Horizon Wimba. This program includes diverse methods of oral delivery, including, but not limited to, asynchronous and synchronous forums, voice presentations, and oral assessment tools. The students were divided into four groups (A-D) and each group was responsible for doing specific listening and speaking (perception and production) exercises that

included the targeted phones (the Spanish voiced-voiceless stops, the tap & trill) various times per week, for fifteen-twenty minutes each day in order to complete the hour. Due to the nature of this study and paper, I will focus on just one, the audio-visual imitation (imitation with video) exercise that was created using Voice Presentation (see Figure 5). Group D was instructed to do three tasks that required the usage of two videos that appeared on the right hand side of the image. The first task asked the students to listen, view, and imitate the person on the video as close as possible because studies have shown that IMITATION, the act of performing an action as performed by a visual model as exactly as possible, involves the firing of more neurons in different areas of the brain (Gallese et al., 1996), this in turn may help with the longer retention of the sounds. Here, I will not delve into specifics, but another goal of the project is to pin point the exact areas of the brain involved with the hope that it will serve as additional evidence that an accent can be eliminated provided that you use proper training. Feedback was given whenever possible.

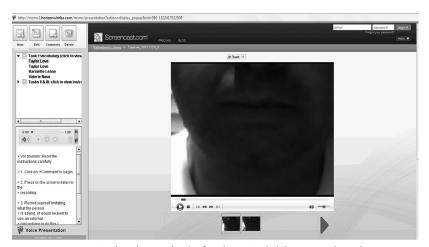


Figure 5. Wimba, the method of audio-visual delivery to the subjects.

5.3 ANALYSIS PROCESS. In order to determine whether the subjects (students) were producing the aforementioned phonemes with a native-like pronunciation (or accent), two assessments were taken: acoustic measures and ratings for accents by native speakers of Spanish. The latter was conducted by native speakers of Spanish (monolinguals from Chile, Colombia, Mexico, Puerto Rico and bilinguals from the Los Angeles area). In order for this to be feasible, pre and post assessment recordings

were digitized and numerically labeled in order to keep the subjects anonymous. The recordings were then uploaded to a server and a private link was given to the raters so that they would have sole access to them. A survey was designed so that it would help the raters evaluate the subjects' accents. Ratings were completed and collected entirely online. This made it possible for the Chileans to do the ratings from their native land, Chile. Detailed instructions on how to rate were given via email.

As for the acoustic measures, the Voice Onset Times for the plosives /p, t, k, b, d, g/ were measured in milliseconds (ms); the rhotics, the tap /r/ [f] & the trill /rr/ [r], on the other hand, were measured by the number of times the tongue had contact (taps) with the alveolar ridge and whether there was a long or short duration when producing such. Any other pertinent acoustic measures (i.e. F2/F3, etc.) will be included with the final results. In order to understand what was being measured and to know for certain that the subjects were producing more L2 phonemes than L1s, I included below brief acoustic descriptions of the Spanish and English plosives and rhotics. For the productions of plosives, the parameters that were created by Cho & Ladefoged (2000) were used. Their parameters were taken from utterances produced by native speakers of English and Spanish.

According to Cho & Ladefoged (2000) English voiceless stops have substantial delay between release and the onset of laryngeal vibration, the VOTs are 30ms or longer. English voiced stops in initial position are generally not pre-voiced, but released simultaneously with the onset of the voicing, for a VOT of approximately zero. Spanish voiceless stops, on the other hand, are produced with near simultaneous release and onset of laryngeal vibration. The VOTs are approximately zero or a few milliseconds 0-10ms. These are very similar to the English voiced stops. The Spanish voiced stops, in contrast, are pre-voiced such that the onset of voicing precedes the release by 40ms or more. In other words, their VOTs are negative and are less than 40ms. When the speaker uses the wrong parameter, we have what is called ACCENTED SPEECH. This is one of the reasons why we say that a person has an accent.

	Voiced Stops /b, d, g/	Voiceless Stops /p, t, k/ 30 ms or longer approx. Zero to 10 ms	
English Avg.VOTs	approx. zero		
Spanish Avg. VOTs	< -40 ms		

TABLE I. Parameters set by Cho & Ladefoged (2000)

Once again, the goal for this ongoing study is to determine whether L2 learners are able to produce tokens with the correct parameters as a result of completing the specific imitation tasks that were created. In the next section, only the preliminary results for the rhotics will be given in addition to a few examples of the productions.

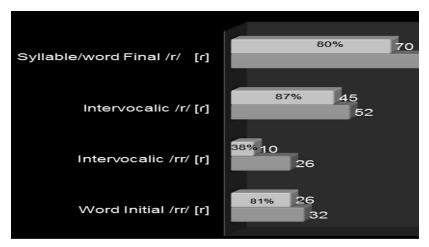
6. Preliminary results and discussion

6.1 Rhotics. When dealing with the productions of the Spanish trill /rr/, we are looking for 2 to 5 closures, taps to the alveolar ridge (Recasess and Pallares, 1999). The most common productions are 3 taps as indicated by Quilis (1993), Blecua (2001) & Sole (2002). These can be seen, once again, in figure 4. The results presented here are those productions that native speakers of Spanish and English were asked to rate by using an excel sheet and a simple rating rubrics.

Rating Rubrics			
0=don't know (write what you think you hear)			
1=Spanish 'rr' as in caRRo			
2=Spanish 'r' as in 'caRo'			
3=English 'r' as in Rose			
4= [w] As in 'wabbit'			
5= Any other production of /r/ or /rr/			

Table 2. Rubrics used to rate productions of rhotics

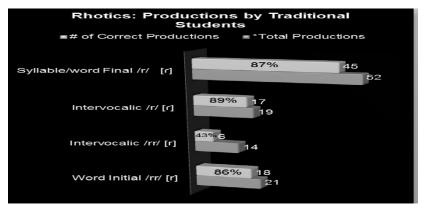
For the results below, it is important to indicate that the tokens contained the same phones within the same environment, but not all were produced by the same individuals. For this reason, the number of total actual productions for each group was counted and then divided by the number of correct ones. These graphs depict some of the results. A statistical analysis will be done at a later time in order to determine whether there are any significant differences between both teaching environments, traditional and online.



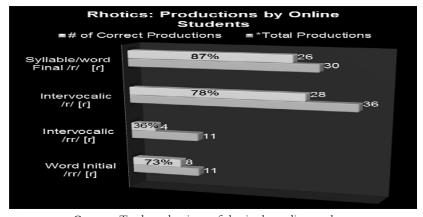
Graph 1. Total productions of rhotics

GRAPH I depicts the total (clear) production of Spanish Rhotics by online and traditional students. When dealing with syllable or word-final /r/ as in *castigar* 'to punish', *decir* 'to say', *puerta 'door*', we find that 80% produced them correctly, well above average; 87% for intervocalic /r/ as in *caro* 'expensive' and *pero* 'but'; 38% for the intervocalic /rr/ (the most difficult) as in *carro* 'car', *perro* 'dog' and *carrera* 'race/career'; 81% for word-initial /rr/ as in *rato* 'a while, little while, short time', *ruido* 'noise' and *regar* 'to water', to name a few.

When we break them down by groups, traditional vs. online as seen in Graphs 2 & 3, we find that the differences among them are insignificant. Both groups, coincidentally produced 87% of the syllable/word final /r/ correctly, they had the most difficulty with the intervocalic /rr/ and for the other two (word initial /rr/ and intervocalic /r/), the productions were average or slightly above average productions.



GRAPH 2. Total productions of rhotics by traditional students



Graph 3. Total productions of rhotics by online students

It is worth noting that when dealing with phrase or word final position /r/, the higher the vowel, the easier it was for the subjects to produce native or near-native like tokens. If we want L2 students to produce all instances of the rhotics in any environment correctly, it would be a good idea to begin with those environments that they po know how to articulate well (i.e. word initial /rr/) and then transfer that knowledge to a more complex environment (i.e. intervocalic /rr/). This can be done by using the imitation task that was depicted in section 5. The task helps the subjects become more aware of what they are doing, having them pay closer attention to those particular articulatory cues that they will eventually transfer to the complex environment.

As for the acoustic measures that will be analyzed, in order to have visual evidence of whether the subjects are attaining native-like pronunciation of the rhotics, we look at the number of taps and the duration of the production. Both can be seen within the pink selection in Figure 6. This one depicts the pre and post productions of the intervocalic trill as in / perro/. As can be seen, prior to the training tasks, the subject was unable to produce a full trill. There were no visible taps and the duration was much shorter, 19ms. Evidence of its correct production is seen in the after spectrogram, where there are at least two taps and duration is much longer, 114ms.

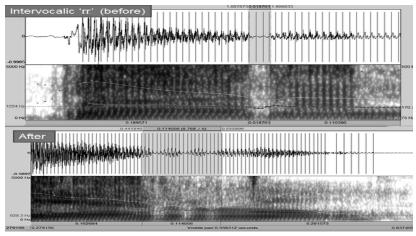


Figure 6. This figure illustrates the pre and post production of intervocalic trill.

FIGURES 7 and 8 below are examples of a subject's productions of the minimal pairs, PERO 'but' and PERRO 'dog'. Prior to the training, the subject was unable to make the distinction. The tap was produced for both. But, as can be seen, there is a clear distinction between both. The duration for the intervocalic trill, FIGURE 7, is a lot longer than the one for intervocalic tap, FIGURE 8.

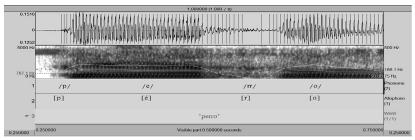


Figure 7. Example of intervocalic trill /rr/ [r]

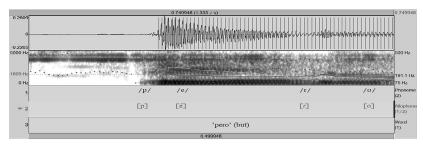


FIGURE 8. Example of intervocalic tap /r/ [f]

7. Conclusion. The purpose of this ongoing study is to encourage L2 students and instructors, and those in the field of language acquisition, to place more emphasis on the accuracy of pronunciation as this one is the key to the success of many of our L2 learners. In order to guide students to a more native-like production, it is highly recommended that AUDIO-VISUAL IMITATIONAL tasks be implemented. These will help L2 learners become aware of the specific acoustic information or cues found within the articulatory gestures that will definitely help with the ACCURATE L2 production.

It is worth noting that this paper by no means is an exhaustive account of all of the neurological processes that need to take place in order for an L2 learner to attain a more native-like accent. Current research being conducted by the author will shed some light on the underlying mechanisms involved in imitation and L2 learning.

Notes

- 1. To the best of my knowledge this term has not been used in the context that is being presented here.
- 2. This term was taken from the ACTFL 2012 guide. In this paper, it will serve as a determinant of the accentedness of speech.
- 3. This idea or connection is comparable to the one that the internationally known otolaryngologist Dr. Alfred Tomatis (1920–2001) discovered back in the midtwentieth century, the connection between the ear and the voice. He emphasized that we can only produce what the ear can hear. He originally discovered it in 1947, but it was not until ten years later that he made it public.

REFERENCES

Barry, W. 1989. Perception and production of English vowels by German learners: instrumental-phonetic support in language teaching. *Phonetica* 46. 155-168.

- Blecua, Beatriz. 2001. Las vibrantes del español: Manifestaciones acústicas y procesos fonéticos. Barcelona, Spain: Universidad Autónoma de Barcelona dissertation.
- Brown, Cynthia. 2000. The interrelation between speech perception and phonological acquisition from infant to adult. *Second language acquisition and linguistic theory*, ed. by J. Archibald, 4–63. Malden, MA: Blackwell.
- Brown, H. D. 2001. Teaching by principles (2nd ed.). White Plains, NY: Pearson.
- CARAMAZZA, A., YENI-KOMSHIAN, G. H., ZURIF, E. B., and CARBONE, E. 1973. The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America* 54. 421–428.
- Cно, T. and Ladefoged, P. 1999. Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics* 27. 207–229.
- Cho, T., Jun, S.-A., and Ladefoged, P. 2000. An acoustic and aerodynamic study of consonants in Cheju, *Speech Sciences* 7. 109-142.
- FLEGE, JAMES EMIL. 1992a. The intelligibility of English vowels spoken by British and Dutch talkers. *Intelligibility in speech disorders: Theory, measurement, and management*, ed. by R. D. Kent, 157–232. Philadelphia: Benjamins.
- FLEGE, JAMES EMIL. 1992b. Speech learning in a second language. *Phonological development: models, research, implications*, ed. by C.A. Ferguson, L. Menn, and C. Stoel-Gammon, 565-604. Timonium, MD: York.
- FLEGE, JAMES EMIL. 1995. Second-language speech learning: Theory, findings, and problems. Speech Perception and Linguistic Experience: Issues in Cross-language Research, ed. by W. Strange, 229 273. Timonium, MD: York Press.
- Gallese V., Fadiga L., Fogassi L., Rizzolatti G. 1996. Action recognition in the premotor cortex. *Brain* 119. 593-609.
- Kumaravadivelu, B. 2003. Beyond methods: Macrostrategies for language teaching. New Haven, CT:Yale University Press.
- Kumaravadivelu, B. 2005. Understanding language teaching: From method to postmethod. Mahwah, NJ: Lawrence Erlbaum.
- Lisker, L. and Abramson, A.S. 1964. A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20. 384-422.
- MAJOR, R. C. & KIM E. 1996. The Similarity Differential Rate Hypothesis. *Language Learning* 46.465–496.
- Meltzoff, A. N. & Moore, M. K. 1999. A new foundation for cognitive development in infancy: The birth of the representational infant. *Conceptual development: Piaget's legacy*, ed. by E.K. Scholnick, K. Nelson, S. Gelman, & P. H. Miller, 53–78. Mahwah, NJ: Erlbaum Press.
- Munro, M. J. & Derwing, T. M. 1995a. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning* 45(1). 73–97.
- Munro, Murray J. and Tracey M. Derwing. 1995b. Processing Time, Accent, and Comprehensibility in the Perception of Native and Foreign-Accented Speech. *Language and Speech* 38(3). 289–306.
- Quilis, Antonio. 1993. Tratado de fonología y fonética españolas. Madrid: Gredos.

- Recasens, Daniel and Maria Dolors Pallarès. 1999. A study of /P/ and /r/ in the light of the «DAC» coarticulation model. *Journal of Phonetics* 27.143–69.
- Solé, Maria-Josep. 2002. Aerodynamic characteristics of trills and phonological patterning. *Journal of Phonetics* 30.655–88.
- Solé, Maria-Josep, John Ohala, and Goangshiuan Ying. 1998. Aerodynamic characteristics of trills. *Proceedings of the 16th. International congress of Acoustics and the 135th. Meeting of the Acoustical Society of America*, 2923–2924.
- Ткиветzкоv, N.S. 1939. Grundzüge der Phonologie, Travaux du Cercle Linguistique de Prague, 7; Prague: n.p. Reprint, Gottingen: Vandenhoeck & Ruprecht 1958.
 Translated into English by Christiane A.M. Baltaxe as *Principles of phonology*. Berkeley: University of California Press, 1969. French translation by J. Cantineau: Principes de phonologie.