Explicit Phoneme Instruction: A Second Look at a Classroom Study

Mark DONNELLAN
James C. JENSEN
Michael GRECO

This article updates and strengthens Jensen's (2010) classroom listening study. The study tested the success of the lab-based training studies that improved the perception of difficult nonnative phonemes. Two research questions were asked: 1) To what extent can the success of lab-based perceptual training studies be replicated in an EFL classroom? 2) How will students react to the training? This article re-examines the data from the original study using more rigorous statistical analysis. The results of the paired t-test of the participants' pre and post-tests reveal statistically significant improvements in their test scores. The students were asked to respond to a short survey to gage their opinions about the training. The results of the survey indicated that the students valued the training highly.

BACKGROUND

There has been little research on the effectiveness of explicitly teaching English phonology in an English as a Foreign Language (EFL) setting (Couper, 2006). By reporting on a classroom study replicating successful lab-based studies that improved the acquisition of nonnative phonemes, "Explicit Phoneme Instruction: A Classroom Study," (Jensen, 2010), took a step in addressing this gap in the research. The current article re-examines the data and updates the literature of Jensen's article.

Over the last couple of decades, studies have shown that training leads to improvement in the discrimination of nonnative phonemic contrasts (Strange & Dittmann, 1984; Logan, Lively, & Pisoni, 1991; Zhang, Kuhl, Imada, Kotani & Tohkura, 2005). The goal of the training and testing in Jensen's 2010 study was to investigate whether techniques that were effective in previous lab-training studies are also effective in an EFL classroom.

Native Japanese speaking university students were tested and trained on three English minimal pairs that are non-existent in Japanese. Two research questions were asked:

1) To what extent can the success of lab-based perceptual training studies be replicated in an EFL classroom?

2) How will students react to the training?

The effectiveness of the training was measured by pre-training and post-training tests. The training was carried out over a 14-week semester, and the students' reactions to the instruction were gauged by a questionnaire. The results of the testing showed that clear gains in perception were made. The questionnaire showed that learners believed instruction on individual phonemes to be valuable. The current article will provide a rigorous statistical analysis of Jensen's (2010) data.

The first issue that is addressed is why such training is important. The article will then discuss the best methods for improving students' perceptive abilities. Finally, we will describe and discuss the classroom study.

LITERATURE REVIEW

There are a number of arguments supporting phoneme acquisition training. It can be said that when learning a language, you have to learn to hear the unhearable. In this section we will look at the most compelling reasons for explicit phoneme instruction in an EFL context.

The relationship between speech perception and production

One argument supporting phoneme acquisition training is the association between perception and production. Although L2 research has shown inconsistent results regarding this relationship, most studies indicate that accurate perception is a prerequisite for good production. A slightly dated review of the relevant studies can be found in Llisterri (1995). Along the same lines, and more recent, is Escudero (2005). After considering studies that may contradict the fact that L2 perception develops before production, Escudero concludes that such studies had "experimental shortcomings," and that from the "weight of the evidence, it can be concluded that perception develops first and needs to be in place before production development can occur" (p. 3).

In a relevant study, Bradlow, Akahane-Yamada, Pisoni, and Tohkura (1999) trained eleven native Japanese speakers on the perception of English /r - l/ over the course of three months. The results demonstrate that even though the speakers were only trained on perception, their production of the /r/ and /l/ phonemes improved dramatically. The researchers suggest, "the underlying mechanism that facilitates the transfer and retention of learning in the perceptual domain to the production domain is due to training-induced modifications to a common mental representation that underlies both speech perception and speech production" (p. 983).

Phoneme acquisition improves other language skills

Another reason supporting explicit training in phonemic acquisition is that phoneme acquisition "bootstraps" other language skills. A study, "Phonetic training makes word learning easier" (Perfors & Dunbar, 2010), indicates that training on novel phonetic contrasts improves word learning. Werker & Yeung (2005) show that phoneme perception abilities are related to later word learning in young children. Additionally, Zhang and Cheng (2011) summarize the research on the role

phonemic knowledge plays in language learning and claim, “The empirical data provide support for both longitudinal and cross-sectional associations between phonetic perception and language skills” (p. 11).

To understand how the ability to distinguish phonemes aids in word learning, the authors refer to yet pertinent article, “The Phonological Loop as a Language Learning Device” (Baddeley, Gathercole, & Papagno, 1998). In this article the phonological loop is understood to be an element of short-term memory’s executive control mechanism. It stores speech sounds in their temporal order. A detailed explanation of this mechanism is beyond the scope of this paper, but a pertinent quote points out the phonological loop’s importance to learning a second language.

> …[There is] evidence that the phonological loop plays a crucial role in learning the novel phonological forms of new words. The authors propose that the primary purpose for which the phonological loop evolved is to store unfamiliar sound patterns while more permanent memory records are being constructed. (p.1)

It would follow that if the input, the sound patterns, were not clearly distinguished, the subsequent memory functions would suffer.

Non-native to non-native communication

Perhaps the most compelling reason to teach phoneme acquisition is the work of Jennifer Jenkins (2000). Jenkins, who has compiled what is perhaps the largest body of data concerning communication between two non-native speakers (NNS), concluded that mistakes at the segmental level, that is the phonemic level, were the biggest source of communication problems. More precisely, when both speakers are NNS, errors on the phonemic level cause more breakdowns in communication than errors with grammar, lexicon and notably, the supra-segmental level. The latter point goes against what mainstream ESL theory maintains, draws attention to the differences between ESL and EFL, and calls for a reassessment of priorities. Given the status of English as the global language and the possibility that students are as likely to use English with other NNSs as they are NSs, improving phoneme perception should be recognized as a basic skill that needs refining and practice.

Learner Expectations

Lastly, teaching pronunciation fits into what the students, particularly in Asia, may expect. In societies where teacher-centered classes are the norm, teacher-directed practice might be familiar, and well understood in terms of required learning goals.

HOW TO TEACH SEGMENTALS

To describe the most effective methods for improving phoneme acquisition one needs to understand a learner’s phoneme architecture prior to any L2 experience. Jensen (2010) used the Native Language Magnet Theory (NLM) as

described in Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola and Nelson (2008). NLM conceptualizes an L2 learner's initial state and describes the difficulties they are faced with when hearing the sounds of the L2. NLM has also proven to be a successful model for perceptual training studies (see Zhang et al., 2005).

NLM contends that language experience "warps" perception. Experience, this argument holds, alters perception. "No speaker of any language," writes Kuhl (2000), "perceives acoustic reality; in each case, perception is altered in the service of language" (p. 11853). This statement is one of the most interesting organizing principles to emerge from her research and explains how infants organize input to recognize similarities and form categories called a sound map.

NLM claims that an infant's creation of mental sound maps "commits" neural structure in some way, and this "neural commitment to a learned structure interferes with the processing of information that does not conform to the learned pattern" (Kuhl, Tsao, Liu & de Boer, 2001, p. 161). Early learning, the argument goes, has committed neurons to a particular configuration, and the network has reached a point at which it can no longer revert to its original plasticity.

This L1 entrenchment causes sounds close to the prototype /i/, for example, to be heard as an /i/. This is what Kuhl (1991) calls the *perceptual magnet effect.* Once a sound category exists in memory, according to this theory, "it functions like a perceptual magnet for other sounds in the category" (Kuhl, 2000, p. 11853). That is, the prototype attracts sounds that are similar so that they sound like the prototype itself. This is why Japanese, who do not have the prototype of the vowel of "bit" mapped in memory, tend to hear it as the vowel in "beat" which they do have mapped.

This commitment to one's native language comes early in life. Research has documented a perceptual transition from "language-general" to "language-specific" perception during the first year of life (Werker & Tees, 1983). Furthermore, studies have illustrated the profound effects of first-language interference in speech perception. Second or foreign language learners cannot hear all the sounds of the language they are learning (Kuhl, 2000). In his influential book, *Biological Foundations of Language,* Lenneberg (1967) proposed that language is constrained by biology. He proposed that language could be acquired only during a critical period, which lasts from birth until the onset of puberty. However, Kuhl and her colleagues, in collaboration with NTT Labs in Tokyo, proved neural plasticity remains well past puberty. The apparent decline in our ability to learn languages, they showed, is not a biological event. Other factors are involved, which they demonstrated by improving the ability of Japanese university students to distinguish /r/ from /l/ (Zhang et al., 2005). The salient point they made is that learners have to be able to hear the sounds to learn them. In the next section, we will discuss the method that has proven most successful for phoneme acquisition.

Signal Enhancement

Humans lose the ability to hear some sounds of foreign languages at an early age (Werker & Tees, 1983), and this perceptual shift necessitates explicit

training. L2 learners cannot distinguish the sounds so it is extremely doubtful they will pick these sounds up from natural communication. Explicit instruction is needed. McCandliss, Fiez, Protopapas, Conway and McClelland (2002) demonstrated the problem experimentally. In their study, subjects who received exposure to contrastive phonemes they could not distinguish failed to benefit from training, even with hundreds of exposures. Subjects who received exposure to stimuli that they could discriminate (i.e., exaggerated stimuli) "showed considerable gains in both identification and discrimination" (p. 93). In short, they had to hear the difference before they could learn it. The way to accomplish this was simple: the input was contrasted and enriched, that is exaggerated, so the learners could distinguish them. Zhang and Cheng (2011) say, "… importantly, speech modifications and acoustic enhancements, … achieves some remarkable success in training second-language learners and treating children with language disabilities" (p. 11).

There are studies showing different results that seem to contradict McCandliss et al. (2002), such as Flege, Bohn and Jang (1997). Flege, Bohn and Jang, for example, demonstrated that length of exposure to non-native contrasts can implicitly influence perception abilities. These studies, however, were conducted in ESL settings, in countries where English is widely spoken. The amount of input the subjects received is far greater than what most EFL students receive. For those in a country where English is not widely used, the input needs to be enhanced so the learners can distinguish them.

High Variability

Another point confirmed by previous studies (for a review see Logan, Lively & Pisoni, 1991) is that exposure to multiple speakers, "high variability," as it is called, seems to be an effective way to ensure that the learning generalizes to never-before heard speakers (Zhang, et al., 2005). Variability in speech tokens is thought to be helpful because it highlights the context of the acoustic parameters and trains the listeners to extract the key acoustic cues. The study described in this article used CDs and DVDs to accomplish the high variability threshold.

THE STUDY

The subjects of this study were 75 Japanese university students. They were all second-year English majors. Their TOFEL scores ranged from 420 to 460. They comprised three classes. The classes taught all four-language skills but focused on speaking and listening. The classes met twice a week.

Method

*Pre-test***.** Two contrasts were tested: /b – v/, and the *a* sound in cat /æ/ and the *u* sound in cut /ʌ/. In the pre-test, students were given handouts with word-pairs containing the contrasts (see Appendix A) in two columns, one column under A and the other column under B. The tester then pronounced one of the words and the students chose either A or B depending on which column the word was listed under. The tester spoke American English and covered his mouth so as not to give visual hints. The test was given the first day of class.

Training. The class met Tuesdays and Fridays. Almost every Tuesday this

exercise was repeated using different contrasts. The correct result would be given, the pronunciation practiced—with emphasis—to make the contrast salient. The students would then work in pairs, one student saying the words and the other writing A or B. They then changed roles. This gave additional listening practice as well as speaking practice. This took ten to fifteen minutes. On Fridays other activities for pronunciation practice including tongue twisters and dictations were utilized.

Much of the exposure the students subsequently received came during more-communicative activities. That is to say, /b - v/, for example, would be exaggerated (not always of course) during normal classroom administration. While giving instructions, explaining meaning-based activities, or during any other classroom discussion, the contrast might be enhanced, which can be taken to the absurd for comical effect. When practicing supra-segmental features, the students were reminded of the difference in articulating the phonemes. In teaching reduced sounds, how words blend and other supra-segmental features, the teacher might use /v/ and exaggerate the sound to make it salient and to point out the contrast to /b/. Clearly, blending can be practiced with attention given to the segmental features.

The contrasts also became part of the teacher's error correction repertoire. Applying the tenets of a form-focused approach (Long & Robinson, 1998), explicit error correction was considered appropriate in meaning-based contexts. Given the appropriate circumstances, an error was indicated. Since self-correction is the ultimate goal of feedback, the teacher might repeat the incorrect sentence with some emphasis on the error (Did you play *bolleyball*?).

Likewise, facial expressions and /or gestures can be used. In the /b - v/ example, the author uses a gesture established early in the term, he bites his lower lip. This not only indicates that an error was made, but also pinpoints it. The effectiveness of using gestures to correct a pronunciation error was confirmed when a student used the lip-biting gesture to point out a pronunciation error her partner committed during a purely communicative task.

The high variability called for by the successful training studies discussed earlier was reached by using a DVD [*b4* from aka-kara English (aka-kara.com)] produced specifically to give listeners exposure to contrastive minimal pairs. On this DVD a number of speakers pronounce different examples of a given contrast. The DVD was used every other week, took no more than six minutes and was easily worked into the class.

Post-training test. The post-training test was identical to the pre-training test. The same contrasts were used and spoken by the same tester.

RESULTS

In order to address the first research question, the pre- and post-test scores were submitted to statistical analysis using SPSS.

TABLE 1
Paired sample tests, minimal pairs post-test - pre-test scores (N = 75)

	Paired Differences					t	df	Sig. (2-tailed)
				95% Confidence Interval of the difference				
	Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	.20270	.54847	.06376	.07563	.32977	3.179	74	.002
Pair 2	.35135	.55966	.06506	.22169	.48101	5.401	74	.000
Pair 3	.33784	.60312	.07011	.19811	.47757	4.819	74	.000
Pair 4	.22973	.45473	.05286	.12438	.33508	4.346	74	.000
Pair 5	.14865	.39457	.04587	.05723	.24006	3.241	74	.002
Pair 6	.02703	.64048	.07445	-.12136	.17541	.363	74	.718
Pair 7	.20270	.54847	.06376	.07563	.32977	3.179	74	.002
Pair 8	.04054	.50854	.05912	-.07728	.15836	.686	74	.495
Pair 9	-.02703	.64048	.07445	-.17541	.12136	-.363	74	.718
Pair 10	.37838	.54150	.06295	.25292	.50383	6.011	74	.000
Pair 11	.41892	.59695	.06939	.28062	.55722	6.037	74	.000
Pair 12	.14865	.58946	.06852	.01208	.28522	2.169	74	.033

The results of paired t-tests showed statistically significant increases in correct student responses for 8 of the minimal pairs (drug - drag; much - match; butter - batter; base - vase; a bit - avid; vines - bines; best - vest; curb - curve; $p < .005$ (two tailed)). Three other minimal pairs (stuff - staff; run - ran; cat - cut) showed increases, but these were not statistically significant. One minimal pair, hat - hut, showed a decrease. We will give possible explanations for these results in the Discussion section below.

TABLE 2
Paired sample statistics (N = 75)

	Mean	Std. Deviation	Std. Error Mean
Overall Pre-test	7.6267	1.85103	.21374
Overall Post-test	10.2533	1.76370	.20365

TABLE 3
Paired sample tests, overall post-test - pre-test scores (N = 75)

	Paired Differences					t	df	Sig. (2-tailed)
				95% Confidence Interval of the difference				
	Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Overall	2.62667	2.16716	.25024	2.12805	3.12528	10.497	74	.000

In relation to the overall test scores, a paired-sample t-test was conducted to evaluate the impact of the minimal pairs exercises on students' overall pre- and post-test scores. There was a statistically significant increase from the students' pre-

test scores (M = 7.6267, SD = 1.85103) to the post-test scores (M = 10.2533, SD = 1.76370), t (74) = 10.49, p < .0005 (two tailed). The mean increase in test scores was 2.62667 with a 95% confidence interval ranging from 2.12805 to 3.12528. The eta squared statistic (.598) indicated a large effect size.

The second research question of the study aimed to investigate students' opinions regarding the training. A four-item questionnaire was designed based on Eccles and Wigfield's (1995) model of task values. According to Eccles and Wigfield, the value of a task is made up of its cost in terms of time and effort, its attainment value, its intrinsic value, and its extrinsic value utility. In general terms, the value of explicit phoneme instruction was judged by asking the students their overall feelings (cost), whether they enjoyed it (intrinsic value), whether the instruction was helpful (attainment value), and whether they thought it important (extrinsic utility value). Students responded on a 5-point Likert scale with 1 being strongly disagree and 5 being strongly agree.

TABLE 4
Descriptive statistics (N:75)

	Median	Mode	Range	Inter-quartile range (IQR)
The exercises were enjoyable.	5	5	3	1
The exercises were helpful.	5	5	4	1
The exercises were important.	5	5	4	1
I had a good overall impression of the exercises.	5	5	2	1

TABLE 5
Summary of responses (N:75)

		Strongly disagree	Somewhat disagree	Neutral	Somewhat agree	Strongly agree
The exercises were enjoyable.	#	0	2	11	20	42
	%	0	2.6%	14.7%	26.7%	56%
The exercises were helpful.	#	1	2	4	24	44
	%	1.3%	2.6%	5.3%	32%	58.7%
The exercises were important.	#	1	1	3	22	48
	%	1.3%	1.3%	4%	29.3%	64%
I had a good overall impression of the exercises.	#	0	0	6	26	43
	%	0	0	8%	34.6%	57.3%

All fours questionnaire items yielded similarly positive responses to the minimal pair exercises, with the majority of students indicating that they found the exercises to be extremely important (Mdn = 5, IQR = 1), helpful (Mdn = 5, IQR = 1) and enjoyable (Mdn = 5, IQR = 1), and that they had a very high overall impression of the exercises (Mdn = 5, IQR = 1).

DISCUSSION

Many training studies for difficult L2 contrasts have shown that structured, laboratory training successfully helps L2 learners on such difficult distinctions. (See Iverson, Hazan & Bannister, 2007 for a discussion). The aim of the present study

was to find out if these methods are as equally effective in an EFL classroom. The results of this study have shown that they are. From the data it can be stated that the successes of other phoneme acquisition training studies was replicated in an EFL classroom setting. It can also be stated that the students found the instruction valuable.

Phoneme perception is complex, due to the properties of individual sounds and their position in relationship to other sounds. Issues like the interactions between vowels and consonants, syllable structure, native language interference and other nuances, play a role in perception difficulties (see Zhang & Cheng, 2011). This paper does not address these complexities. Our aim was not to be explanatory in this regard. That said, a few words about our results are called for.

It is not surprising the minimal pairs formed with consonants were the most successfully acquired. Research findings have shown that consonants have an advantage over vowels in terms of accessibility and learnability (Nam, Goldstein & Saltzman, 2009). As for why some vowel pairs made statistically significant gains and others did not, it may have been a function of the consonants they interacted with. Strange, Akahane-Yamada, Kubo, Trent and Nishi (2001), demonstrated that the acquisition of vowel sounds is often a function of the consonantal context. In this study, the vowels that were distinguished were juxtaposed with voiced consonants. The vowels that showed an increase in correct responses that were not statistically significant, were juxtaposed with the unvoiced consonants, /r/ and /n/. Finally, the vowel pair that actually showed a decrease in correct responses followed /h/, a phoneme with different realizations in Japanese. In short, a few minutes of exposure over the course of 14 weeks was sufficient exposure to train the students to distinguish the consonants, but insufficient to accomplish the acquisition of the vowels in all phonemic contexts.

Another limitation of this study is the number of tokens. A larger set of minimal pairs needs to be incorporated; a set that uses twelve pairs is too small. More pairs that place the contrasting phonemes into more phonetic contexts would be informative and might enhance the possibility of teasing the above-mentioned nuances from the data. Also, recent research suggests that using a large set of tokens may be more effective for long term retention (Nishi & Kewley-Port, 2007). Future studies should keep this in mind.

The second research question dealt with the students' attitudes toward pronunciation exercises that focus on phoneme acquisition. The result from the questionnaire shows these kinds of activities are welcomed and valued by students. The students were asked to rank different elements of the training from 1 to 5 with 5 being the highest. If 4 and 5 are considered positive reactions nearly 90% of all responses were positive. If 1 and 2 can be considered negative responses, there were only 7 in total, just 2%.

One may suspect some element of *subject expectancy* in the overwhelmingly positive result. That may be the case, although every effort was made to control such a variable. The questionnaire included questions about other aspects of the class and it was given on the day of the final speaking test. There were no listening activities; there was no temporal connection between the

questionnaire and the exercises it inquired about.

Lastly, unlike the lab-based studies that informed it, this study was affected by the realities of a classroom. The trainer's voice quality, for example, may have been inconsistent. The use of recorded materials would eliminate this random effect.

CONCLUSIONS

Clearly, caution must be exercised when drawing conclusions about a study with the limitations this one has. Nonetheless, the findings reported in this study support the view that explicit phoneme instruction can be effective in an EFL classroom. This study also shows that the students were extremely enthusiastic about this approach. The current article has sharpened the statistical analysis of the original publication adding substantially to the study's validity. This article has also added support from more recent research to bolster its claims. This brings us to the concluding remark.

Some students claimed they had never experienced such instruction. Some did not know that /v/ is a voiced /f/. One might suspect this is because the tenets of the communicative approach have determined that such explicit instruction is a waste of time. It may be in an ESL situation; the authors cannot comment on that. The authors believe, however, that those of us in an EFL situation need to reconsider if and how we teach both listening and pronunciation. Our students do not receive the input students in an ESL situation do. Research that comes from countries where English is the dominant language should be considered in that light. We need to discover what works and does not work in an EFL classroom. More studies like this one would be helpful.

REFERENCES

Baddeley, A. (2010). Long-term and working memory: How do they interact? In L.Bäckman & L. Nyberg (Eds.), *Memory, aging and the brain: A Festschrift in honour of Lars-Göran Nilsson.* (pp. 7-23). New York, NY: Psychology Press.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The Phonological Loop as a Language Learning Device. *Psychological Review*, *105*(1), 158-173.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics, 61*(5). Retrieved from: http://babel.ling.northwestern.edu/~abradlow/bradlow-etal-rl5-p&p99.pdf

Couper, G. (2006). The short and long-term effects of pronunciation instruction. *Prospect, 21*(1), 46-66. Retrieved from: http://www. Amepre.mq.Edu.au./prospect journal/volume 21 no.1 Couper.pdf

Eccles, J. S. & Wigfield, A. (1995). In the mind of the achiever: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin, 21*, 215-255.

Escudero, P. (2005). Linguistic perception and second Language acquisition: Explaining the attainment of optimal phonological categorization. LOT Dissertation Series 113, Utrecht University.

Flege, J., Bohn, O., & Jang, S. (1997). Effects of experience on non-native speaker's production and perception of English vowels. *Journal of Phonetics, 25*, 437-470. Retrieved from: http://jimflege.com/files/Flege_Bohn_effect_experience_JP_1997.pdf

Golestani, N., Molko, N., Dehaene, S., Le Bihan, D., & Pallier, C. (2007). Brain structure predicts the learning of foreign speech sounds. *Cerebral Cortex, 17,* 575–582. doi: 10.1093/cercor/bhk001

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/ l / to Japanese adults. *The Journal of the Acoustical Society of America, 118*, 3267–3278. DOI: 10.1121/1.2062307

Jenkins, J. (2000). *The Phonology of English as an International Language.* Oxford: Oxford University Press.

Jensen, J. (2010). Explicit phoneme instruction: A classroom study. *Kinki University English Journal,* (5)*,* 81-94. Retrieved from: http://kurepo.clib.kindai.ac.jp/modules/xoonips/ listitem.php?index_id=6911

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics, 50*, 93-107. DOI: 10.3758/BF03212211

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Science, 24*, 11850-11857. Retrieved from: http://ilabs.washington.edu/kuhl/pdf/Kuhl_2000.pdf

Kuhl, P. K., Tsao, F.-M., Liu, H.-M., Zhang, Y., & de Boer, B. (2001). Language/ Culture/Mind/Brain: Progress at the margins between disciplines. In A.

Domasio et al. (Eds.), *Unity of knowledge: The convergence of natural and human science* (136-174). New York: The New York Academy of Sciences. Retrieved from: http://ilabs.washington.edu/kuhl/pdf/2001_Kuhl_etal.pdf

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophic Transactions of the Royal Society B, 369,* 979-1000. DOI:10.1098/rstb.2007.2154

Lenneberg, E. H. (1967). *The biological foundations of language*. New York: Wiley.

Llisterri, J. (1995). *Relationships between speech production and speech perception in a second language*. Retrieved from: http://liceu.uab.es/~joaquim/publicacions/Prod_Percep.html

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/ : A first report. *Journal of the Acoustical Society of America, 89*, 874-886. DOI: 10.3758/BF03206911

Long, M., & Robinson, P., (1998). Focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition.* New York: Cambridge University Press.

McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., & McClelland, J. A. (2002). Success and failure in teaching the [r]–[l] contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, & Behavioral Neuroscience, 2*(2), 89-108.Retrieved June 10, 2008, from http://www.cnbc.cmu.edu/~jlm/papers/McCandlissETAL02.pdf

Nam, H., Goldstein, L., and Saltzman, E. (2009). Self-organization of syllable structure: a coupled oscillator model. In F. Pellegrino, E. Marsico, and I. Chitoran (Eds.), *Approaches to Phonological Complexity* (pp. 299–328). Berlin: Mouton de Gruyter.

Nishi, K., Kewley-Port, D. (2007). Training Japanese listeners to perceive American vowels: Influence of training sets. *Journal of Speech, Language, and Hearing Research, 50*, 1496-1509. DOI:10.1044/1092-4388(2007/103)

Perfors, A., & Dunbar, D. (2010). Phonetic training makes work learning easier. *Proceedings of the 32nd Annual Conference of the Cognitive Science Society.* Portland, OR. Retrieved from: http://mindmodeling.org/cogsci2010/papers/0418/paper0418.pdf

Strange, W., & Dittmann, S. (1984). Effects of discrimination training on the perception of r–l/ by Japanese adults learning English. *Perception and Psychophysics, 36*, 131-145. DOI: 10.3758/BF03202673

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., & Nishi, K. (2001). Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners. *The Journal of the Acoustical Society of America*, *109*, 1691–1704.

Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology, 37*(2), 278-286. doi:10.1016/S0163-6383(84)80022-3

Werker J.F. & Yeung, H.H. (2005). Infant speech perception bootstraps word learning. *Trends in Cognitive Science, 9*(11). DOI: 10.1016/j.tics.2005.09.003

Zhang, Y., & Cheng, B. (2011). Brain plasticity and phonetic training for English-as-a-second language learners. In D.J. Alonso (Ed). *English as a Second Language* (pp. 1-50). Hauppauge, New York: Nova Science Publishers.

Zhang, Y., Kuhl, P., Imada, T., Kotani, M, & Tohkura, Y. (2005). Effects of language experience: Neural commitment to language-specific auditory patterns. *NeuroImage, 26*, 703-720. doi:10.1016/j.neuroimage.2005.02.040

APPENDIX A: List of Minimal Pairs

Pair 1	drug - drag
Pair 2	much - match
Pair 3	butter - batter
Pair 4	base - vase
Pair 5	a bit - avid
Pair 6	stuff - staff
Pair 7	vines - bines
Pair 8	run - ran
Pair 9	hat - hut
Pair 10	best - vest
Pair 11	curb - curve
Pair 12	cat - cut