It Starts With Phonemes

James C. Jensen
Aka-Kara English

Reference Data:
Jensen, J. C. (2013). It starts with phonemes. In N. Sonda & A. Krause (Eds.), *JALT2012 Conference Proceedings*. Tokyo: JALT.

In this paper I discuss the relationship between infant phoneme acquisition studies and teaching in an EFL context. I start with a historical account of the developments in infant language acquisition studies. Three questions are addressed. First, to what extent can second language learners' first language phonemic architecture be modified? Second, what do infant language acquisition studies tell us about the constraints that cause difficulties for such modifications? And third, what methodologies are the most effective for modifying a foreign language learner's native phonemic architecture? I argue that explicit instruction at the segmental level is vital in an EFL setting.

本論文は、乳幼児音素習得研究とEFLコンテキストにおける教育との関連性について論じたものである。まず、乳幼児言語習得研究発展の歴史について論じる。三つの質問が提起される。第一に、どの程度、第二言語学習者の第一言語音素アーキテクチャは修正され得るのか。第二に、乳幼児言語習得研究はそのような修正を困難にする制約に関し我々に何を伝えるのか。そして第三に、どの方法が外国語学習者のネイティブ音素アーキテクチャを修正するのに最も有効であるのか。本論はEFLセッティングにおいて、分節レベルにおける明瞭な指示が重要であると主張する。

To understand the nature of language learning it is crucial to study its earliest stages. The process through which children learn their native language or languages remains a mystery but distinguishing the phonemes, the smallest units of sound, is one of the earliest stages. Research has made huge strides in understanding the process behind this learning and the theoretical implications are extensive, reaching into the foreign language classroom. In this paper I will explore these implications.

I will show that an infant's perception progresses from a language-general to a language-specific state. The task of a native language learner, then, can be characterized as a "mapping" of the L1 phonetic system. In contrast, a nonnative language learner must progress from a system mapped to the sounds of the L1 to one that can be tuned to the L2.

What infant studies tell us about the difficulties faced by L2 learners will be discussed. I will also show how these studies intersect with the Critical Period Hypothesis (CPH) and argue that plasticity remains well into adulthood. Three questions are addressed. First, to what extent can second language learners' first language phonemic architecture be modified? Second, what do infant language acquisition studies tell us about the constraints that cause difficulties for such modifications? And third, what methodologies are the most effective for modifying a

foreign language learner's native phonemic architecture? An argument for explicit phoneme perception training will be made.

Infant Studies

The studies I follow began with Noam Chomsky. In the 1970s researchers set out to evaluate Chomsky's argument that humans' innate constraints for language included specification of a universal grammar and phonetics (Chomsky, 1968). In a landmark experiment in 1971, Eimas (1975) tested infant sensitivity to the sounds of speech and showed that 1-month old babies could differentiate the sounds /pa/ and /ba/.

Eimas's (1975) work was followed by a series of speech perception studies, which revealed that adults could only distinguish those phonetic contrasts that are used in their native language. Infants, on the other hand, discriminate phonetic contrasts whether or not they are used in the language they are used to hearing. Researchers began to accumulate facts that suggested infants were able to discriminate both native and nonnative contrasts equally well (see Saffran, Werker, & Werner, 2006, for a summary). That is, infants can hear the sounds of any language. They are prepared to learn any language to which they are habitually exposed.

The Work of Janet Werker

Clearly, adults cannot distinguish the phonemes of every language, so when the universal listening ability of infants was discovered the question that arose was at what age humans lose this discriminatory ability. The task of answering this question was taken up by Werker. She initially hypothesized that the ability to discriminate nonnative contrasts is lost at puberty. What she found, however, was that the loss of discrimination occurs between 6 and 12 months of age. After 1 year of age, infants have lost much of their ability to discriminate between sounds that are not important in their native languages (Werker & Tees, 1983). In the years since this initial work, there have been a number of replications and extensions of this finding (see Saffran et al., 2006, for a summary). Moreover brain-imaging studies have supported Werker's claim. Cheour et al. (1998) confirmed Werker's results using the MEG brain-imaging technique.

Developmental Change and Learning

Werker's claim that a perceptual shift occurred during an infant's 1st year was supported by research, but evidence began to suggest that maintenance alone is insufficient to capture the dynamics of infant speech perception. Research began to make it apparent that an infant's development is more complicated than what Eimas's model predicted (Kuhl, Tsao, Liu, Zhang, & de Boer, 2001). Eimas's (1975) model held that humans have an innate neural architecture containing all possible phonetic units. The sounds in the ambient language were perceived and so maintained while those not perceived atrophied. Maye and Weiss (2003), however, claimed, "The process of an infant's developing perception of speech must . . . involve not only paring down of initially discriminable contrasts, but also enhancement of initially difficult contrasts" (p. 508). "The framework that emerges from this research is very different from that held historically. Infants are neither the tabula rasa that Skinner described nor the innate grammarians that Chomsky envisioned" (Kuhl, 2000, p. 11856).

The Native Language Magnet Theory

To conceptualize what is occurring in the development of phonemic representations, I will use Patricia Kuhl's Native Language Magnet (NLM) theory as proposed by Kuhl et al. (2008). This theory posited three aspects to an infant's initial learning.

First, infants demonstrate excellent skills at recognizing patterns in speech. Second, infants exploit statistical properties of language input. And third, NLM claims that language experience "warps" perception. "No speaker of any language" wrote Kuhl (2000), "perceives acoustic reality; in each case, perception is altered in the service of language" (p. 11853). The third point is the most relevant because it leads to an understanding of how infants organize input and form categories, i.e., phonetic representations, which she called a sound "map" (p. 11852).

By the time an infant in an English speaking home is 6-months old it has heard hundreds of thousands of examples of the vowel sound /i/ as in *daddy*, *mommy*, and *baby*. Researchers think that from these thousands of examples, babies develop a type of sound map in their brains that helps them hear the /i/ sound clearly. It can be said that babies create perfect examples of speech sounds with a type of target area around each sound. Once their sound map for /i/ is created, babies can pick out the /i/ from the other sounds they hear. These prototypes have a profound effect on how babies hear speech and how they babble (Kuhl et al, 2001). They "tune" the child's brain for the language around them, so they can hear the different sounds of speech clearly. By the time babies are 6 months old, they have developed a set of speech sound prototypes they can use as building blocks when they begin to put together their own words (Kuhl, 1991).

A Biological Critical Period or Interference

Werker demonstrated that early development entails a shift from a language-general to a language-specific pattern of perception. Kuhl et al. (2001) claimed this shift is the creation of a mental sound map that "commits" neural structure, and this "neural commitment to a learned structure interferes with the processing of information that does not conform to the learned pattern" (p. 161). This commitment interferes with later language learning.

To what extent can L2 learners' perceptual patterns be modified after the initial mapping? The Critical Period Hypothesis (CPH) holds that language-learning ability is reduced after puberty as the result of the loss of neurological plasticity of the brain (Lenneberg, 1967). While a full account of CPH is beyond the scope of this paper, in the eyes of some researchers the original and strictest versions of CPH are too simplistic. The view that is emerging is that there may be different critical periods for different language skills, which change at different ages. Many have come to favor the use of "sensitive period" (Tomblin, Barker, & Hubbs, 2007). A critical period is viewed as a time in development in which experience, or the absence of experience, results in irreversible changes in the brain. Sensitive periods, in contrast, do not necessarily result in a complete irreversible change in the brain (Bruer as cited in Tomblin et al, 2007). Commonly, sensitive periods are defined as a time in development when the organism is particularly responsive to experience. Werker and Tees (2005) went further: "to ensure that we are referring to a window that is more variable in onset and offset than a classic CP [critical period], . . . we will employ the term 'optimal period' (OP)." They then said, "Language involves many different subsystems including semantics, syntax, morphology, and phonology—each likely with its own OP or interrelated set of OPs" (p. 234). The evidence for this, they claim, is "overwhelming" (p. 236).

The commonly observed and widely accepted notion that learning gets harder with age is not in question. What is of interest here is what infant studies say about the fundamental changes in the learning process that occur at a fairly fixed age; that is, whether there is a closing of a biological "window of opportunity." My questions are about the possibility of, and the constraints upon, changing the phonemic mapping that occurs during an infant's 1st year. The question concerning us, then, is not if the ability to learn an L2 declines with age; that is uncontroversial. The question is the degree to which L2 learners can improve.

Kuhl, Conboy, Padden, Nelson, and Pruitt (2005), in a study linking infant speech perception to later language development, claimed that "critical period phenomena arise not from a genetically determined change in learning capacity at a particular age, but from entrenchment, which is the direct outcome of learning" (p. 258). This neural commitment to a learned structure, Kuhl (2000) argued, interferes with the processing of information and "initial learning can alter future learning independent of a strictly timed period" (p. 11855).

McClelland (2002) agreed. Discussing the results of a study he conducted of the English /r/-/l/ contrast as perceived by native Japanese speakers he commented,

> The findings . . . suggest that there is considerable residual plasticity in the phonological systems of Japanese adults. Their failure to learn under normal conditions may reflect not so much a loss of plasticity as a tendency for the mechanisms of learning to maintain strongly established perceptual tendencies. (p. 12)

Kuhl (2000) believed that, "L2 speech learning is probably not a strictly timed developmental experience" (p. 1539) and to show this she conducted a training study with Japanese university students (Zhang et al., 2009). This study is noteworthy because the researchers improved the ability of native Japanese speakers to discriminate the English phonemes /r/-/l/ by using characteristics of infant-directed speech (IDS), the kind of speaking parents and caregivers use when speaking to an infant. The theory informing the study is that circumventing the L1 "neural commitment" requires "enriched" exposure analogous to IDS. This idea of signal enhancement has also been found effective in treating children with language disabilities to improve their language skills (Tallal et al., cited in Zhang et al., 2009).

The study by Zhang et al. (2009) was conducted in collaboration with researchers at Nippon Telephone and Telegraph in Tokyo and looked at whether Japanese listeners could be trained to respond to the /r/-/l/ stimuli as linguistic signals, that is with the left hemisphere of the brain (see Kuhl et al., 2001, for an explanation of brain laterality and speech processing). The Japanese subjects heard /r/ and /l/ in syllables, with high fundamental frequency and extended duration. Listeners also heard many different speakers, and the sounds were presented in different vowel contexts. After 12 hours of training the subjects showed over 20% improvement in discrimination. Also, prestudy and poststudy MEG data revealed that the subjects treated more of the stimuli with the left hemisphere of their brains. This indicates that linguistic processing, as opposed to purely auditory processing, was involved. It can be said, then, that neural plasticity remains well beyond puberty.

Why Teach Phoneme Acquisition?

I used NLM to characterize the initial phonemic mapping that occurs with L1. According to this model, the starting point of the neural system for L2 acquisition is, as Ellis (2006) pointed out, a "tabula replete" (p. 184). It has also been shown that NLM conceives this mapping as a neural commitment that becomes entrenched with age and can interfere with L2 learning. While making L2 acquisition difficult, it has been established that plasticity remains well into adulthood and the sound map can be modified. The obvious questions are: So what? Why should an EFL teacher spend class time on the explicit instruction of the phonemes, or as they are often called, segmentals?

I will make three points in arguing for explicit phoneme instruction: (a) improving L2 learners' ability to perceive phonetic distinctions is important for reasons beyond improving listening; (b) explicit training is the only way to improve EFL student's perceptive abilities because they do not get enough input; and (c) evidence suggests that when both speakers are nonnative English speakers, errors involving phonemes (segmentals)

are as important as those on the suprasegmental level, that is on any level higher than the phoneme level.

Regarding the first point, the results of a recent study, suitably titled *Phonetic Training Makes Word Learning Easier* (Perfors & Dunbar, 2010) indicate that training on novel phonetic contrasts improves word learning. To understand how this might work, one can refer to another paper, *The Phonological Loop as a Language Learning Device* (Baddeley, Gathercole, & Papagno, 1998), in which the phonological loop is understood as being responsible for the temporal maintenance of acoustic and speech-based material in working memory.

How might phoneme perception effect performance on other aspects of language? L2 learners have difficulty processing fluent speech in their L2, which may be due to difficulty in perceiving the phonemes that make up that speech. Difficulties in rapid processing could also lead to difficulties in segmenting words. Also, empirical evidence reveals that knowledge of lower level aspects of language (such as phonological perception) can help in the acquisition of more complex linguistic phenomena (Werker & Yeung, 2005). Recent computational work, moreover, suggests that word learning and phonetic category learning are more effective when occurring simultaneously (Feldman & Griffiths, cited in Perfors & Dunbar, 2010). And lastly, the connection between perception and production must be considered. A review of the studies supporting the argument that the development of L2 perception precedes L2 production can be found in Escudero, (2005). And even though Bradlow, Akahane-Yamada, Pisoni, and Tohkura (1999) only trained Japanese native speakers in perception, their production of /r/ and /l/ improved dramatically.

The second point about the importance of explicit phoneme instruction is that such instruction is the only way EFL learners will ever learn to perceive difficult L2 phonemes. According to NLM, we do not hear the spoken sounds directly but filter them through the phonemic structures of our L1. McCandliss, Fiez, Protopapas, Conway, and McClelland (2002) claimed that initial learning creates a strong tendency to treat nonnative contrastive phonemes as a single phoneme from the listener's L1 and that "this tendency may be self-reinforcing, leading to its maintenance even when it is counterproductive" (p. 185). The "sad irony for an L2 speaker," as Ellis (2006) pointed out, "is that more input simply compounds their error; they dig themselves ever deeper into the hole created and subsequently entrenched by their L1" (p. 185). The sounds have to be enriched, like in IDS, so the learner can hear them.

The third point to be made in support of explicit phoneme training is that the mainstream focus on suprasegmentals may not be entirely valid, especially in EFL situations. Suprasegmentals are the features of pronunciation at any level higher than that of the phoneme: things like stress, rhythm, and intonation. Both Levis (2005) and Neri, Cucchiarinin, and Strik (2006) doubted the reliability of the studies ascribing greater importance to suprasegmentals than to the segmentals, and Jenkins (2000) offered some evidence to support their doubts.

Jenkins (2000) viewed English as a *lingua franca* that plays a role in the lives of hundreds of millions of speakers, whom she terms "non-bilingual English speakers" (NBESs). Jenkins looked into the components of English pronunciation that are essential for successful interaction between NBESs and discovered that phonemes, segmentals, are more important than suprasegmentals. Specifically, errors on the level of phonemes caused more communication problems than errors on levels higher than phonemes. Her claim was that the focus on suprasegmentals is unnecessary when teaching NBESs because these elements of spoken English only come into play when one of the conversation partners is a native speaker. Given the status of English as a global language and the possibility that our students are as likely to use English with other nonnative speakers as with

native ones, some work on improving phoneme perception is more than reasonable.

Implications for the EFL Classroom

I have shown that training EFL students to perceive phonemes is possible and that it is important. I will now discuss how a teacher can apply the research to an EFL classroom. First, students have to hear the sounds to learn them. The importance of contrasting and emphasizing L2 phonemes is self-evident. This can be accomplished with minimal-pair exercises, reading individual words, or reading words in sentences. In my classes, I put two columns of minimal pairs on the board, pronounce one word, and ask the students to write the word I say. After giving the correct answer and some doing some practice, I have students do the same in pairs. Tongue twisters and rhymes can also be used. Students can even be asked to create their own tongue twisters and rhymes. Importantly, these drills should be complemented with details about how the sounds are articulated.

Another important point about phoneme acquisition is that exposure to multiple speakers ("high variability," as it is called) seems to be an effective way to increase perceptual learning. Logan, Lively, and Pisoni (1991) showed that the subjects exposed to numerous speakers improved significantly in minimal-pair identification when compared to subjects who heard only one speaker. To expose students to multiple speakers, teachers can use CDs, DVDs, or Youtube clips or make their own videos.

Lastly, these exercises do not take a lot of class time. A recent study by Kase and Jensen (2013) showed that improvement in student perception can be achieved in only a few minutes of class time per week. Of course, when designing a strategy for teaching pronunciation, it is important to consider the nature of the task in terms of the degree of its difficulty for students and the chances for its successful accomplishment. Teachers should also consider the time available and set their goals accordingly.

Conclusion

I have looked at the studies of infant phoneme acquisition and discussed their implications for teaching EFL. Using the NLM theory as a conceptual model, I discussed how infants create a sound map from ambient language and how this sound map represents a neural commitment that interferes with later language learning. Importantly, I showed that explicit phoneme instruction not only "cascades" to higher language skills, but that such perception is critical when two nonnatives are speaking. Contrast, emphasis, and high variability were discussed as effective methods for modifying a L2 learner's initial L1 phoneme structure.

The point I make in conclusion is that training students to hear difficult L2 sounds is best done in the early stages of L2 teaching. I have shown how the ability to hear the difficult sounds of an L2 decreases with age as a result of tendencies that increase with age. Starting early is of particular importance when it comes to improving perception skills. Finally, improving listening skills is inherently important, but the acquisition of these skills also increases confidence and gives students a sense of accomplishment, traits learners will carry with them through their years of instruction.

Bio Data

James C. Jensen is the founder of Aka-Kara English. <jensen@leto.eonet.ne.jp>

References

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158-173. doi:10.1017/S0142716406210397

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*. 977-985. Retrieved from http://faculty.wcas.northwestern.edu/ann-bradlow/publications/1999/bradlow-etal-rl5-p&p99.pdf

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Naatanen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience, 1*. Available from http://www.nature.com/neuro/journal/v1/n5/abs/nn0998_351.html

Chomsky, N. (1968). *Language and mind: Linguistic contributions to the study of mind (future)*. Harcourt Brace Jovanovich. Retrieved from http://www.marxists.org/reference/subject/philosophy/works/us/chomsky.htm

Eimas, P. D. (1975). Speech perception in early infancy. In L. B. Cohen & P. Salapatek (Eds.), *Infant perception: From sensation to cognition* (pp. 193-231). New York: Academic Press.

Ellis, N. C. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics, 27*, 164-194. doi:10.1093/applin/aml015

Escudero, P. (2005). *Linguistic perception and second language acquisition*. Retrieved from http://www.fon.hum.uva.nl/paola/escudero_PhD-manuscript_final.pdf

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.

Kase, K., & Jensen, J. C. (2013). Can a non-native speaker train students in English phoneme acquisition? *Kinki University Center for Liberal Arts and Foreign Language Education Journal 3*(1), 239-250. Retrieved from http://kurepo.clib.kindai.ac.jp/modules/xoonips/detail.php?id=AA12508620-20121130-0239

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. Retrieved from http://link.springer.com/article/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., 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. Retrieved from http://rstb.royalsocietypublishing.org/content/363/1493/979.full?sid=a193c3b1-a271-41cf-a45f-833aadb57d41

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the 'critical period.' *Language Learning and Development, 1*, 237-264. Retrieved from http://ilabs.washington.edu/kuhl/pdf/kuhl_Conboy_2005.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. R. Damasio, A. Harrington, J. Kagan, B. S. McEwen, H. Moss, & R. Shaikh (Eds.), *Unity of knowledge: The convergence of natural and human science* (pp. 136-174). New York: The New York Academy of Science.

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

Levis, J. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39*, 367-377. doi:10.2307/3588485

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:10.1121/1.1894649

Maye, J., & Weiss, D. J. (2003). Statistical cues facilitate infants' discrimination of difficult phonetic contrasts. In B. Beachley, A. Brown, & F. Conlin (Eds.), *Proceedings of the 27th annual Boston University Conference on Language Development* (pp. 508-518). Somerville, MA: Cascadilla 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 from http://psych.colorado.edu/%7Emunakata/csh/mccandlissetal.pdf

McClelland, J. L., (2002). *How far can you go with Hebbian learning, and when does it lead you astray?* Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary? doi:10.1.1.89.1627

Neri, A., Cucchiarini, C., & Strik, H. (2006). *Selecting segmental errors in non-native Dutch for optimal pronunciation training*. Retrieved from http://academic.research.microsoft.com/Paper/4546109.aspx

Perfors, A., & Dunbar, D. (2010). Phonetic training makes work learning easier. *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. Retrieved from http://health.adelaide.edu.au/psychology/ccs/docs/pubs/2010/PerforsDunbar2010.pdf

Saffran, J. R., Werker, J., & Werner, L. (2006). The infant's auditory world: Hearing, speech, and the beginnings of language. In R. Siegler & D. Kuhn (Eds.), *Handbook of child development* (pp. 58-107). New York: Wiley.

Tomblin, J. B., Barker, B. A., & Hubbs, S. (2007). Developmental constraints on language development in children with cochlear implants. *International Journal of Audiology, 46*, 512-523. Retrieved from http://www.uiowa.edu/~clrc/pdfs/trajectories.pdf

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

Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. Published online in *WileyInterScience*. doi:10.1016/S0163-6383(84)80022-3

Werker, J. F., Yeung, H. H. (2005). Infant perception bootstraps word learning. *Trends in Cognitive Science, 9*, 519-527. doi:10.1016/j.tics.2005.09.003

Zhang, Y., Kuhl, P., Imada, T., Iverson, P., Pruitt, J., Stevens, E., . . . & Nemoto, I. (2009). Neural signatures of phonetic learning in adulthood: A magnetoencephalography study. *NeuroImage*, *46*, 226-240. doi:10.1016/j.neuroimage.2009.01.028