Effects of language experience on organization of vowel sounds

Andrew J. Lotto
Dept. of Psychology and Parmly Hearing Institute, Loyola University Chicago
Keith R. Kluender
&
Lori L. Holt
Dept. of Psychology, University of Wisconsin-Madison

There is a long tradition in laboratory phonetics of describing the interaction of phonetic inventories and constraints of auditory perception and articulatory production. For example, Liljencrants & Lindblom (1972) attempted to predict the inventories of typical vowel systems by appealing to the notion of maximal perceptual distinctiveness. Stevens (1989) explained the makeup of these inventories with reference to stable articulatory-acoustic mappings. This tradition is based on the notion that speech is a system which developed to provide robust, reliable and efficient communication. It is presumed that these desiderata have led to linguistic systems that take advantage of the operating principles of the perceptual and articulatory systems and that the structure of phonetic inventories (both supra- and sub-phonemic) will reflect, in some manner, these operating principles (see e.g. Ohala, 1974; Lindblom 1986; Lindblom & Maddieson, 1988; Diehl & Kluender, 1989 for typical examples of this approach). This is certainly a compelling story as far as it goes. However, it is also important that linguistic systems are learnable. Infants and second-language learners must be able to induce the phonetic inventory (and, of course, other aspects of language) to be competent participants in communication. The operating principles of whatever learning is responsible for the induction of phonetic inventories may also leave a fingerprint on the membership and structure of these inventories.

Recent experimental work has been concerned with the description of processes underlying acquisition of phonetic categories. This paper will discuss the validity of some of this work and speculate about what the fingerprints of learning on phonetic inventories may look like.

X.1 Perceptual-Magnets and Phonetic Mappings

The ability of infants (6-month olds) to respond to phonetic segments in a "language-appropriate" manner in the face of phonemically-irrelevant variation has been well established. For example, Kuhl (1983), utilizing a reinforced head-turn paradigm, demonstrated that infants can respond to changes in phonetic identity while ignoring substantial spectral variation due to changes in talker. These functional-equivalence classes for potentially-discriminable sounds have been referred to as "phonetic categories". The presence of these segment-level effects in the perceptual behavior of pre-verbal infants has been taken by some investigators to be an indication of mental representations for phonetic categories (in some form) as prerequisites for speech communication (e.g. Kuhl, 1993).

Recently, investigations into the structure of these functional mappings along phonemically-relevant acoustic/auditory dimensions have been undertaken. Grieser & Kuhl (1989) trained 6-month olds to turn their head when they perceived a change from a repeating /i/ exemplar to another variant of /i/. That is, the ability of infants to discriminate between instances of a single phoneme was tested (as opposed to Kuhl's earlier work testing discrimination of different phonetic segments). Infants were placed in one of two conditions. For one group a very good /i/ variant served as the repeating standard. Formant-frequency values for this vowel were identical to the mean values obtained from male production measurements by Peterson & Barney (1952). Grieser & Kuhl (1989) reported that this exemplar received the highest "goodness" rating (in terms of being an example of /i/) from adult listeners and is, thus, referred to as the "prototype" (**P**). Infants in the second group heard sounds from a second distribution with a poorer representative of /i/, the "nonprototype" (**NP**), serving as the standard. It is important for all of these vowels to be perceived as members of the phonemic class /i/, because this demonstration is intended to be one of intra-class discrimination. The results revealed a significant difference in discriminability of stimuli in the two conditions. Discriminability of the repeating standard and another variant of the vowel was greater when the standard stimulus was the "non-prototype". In other words, infants generalized (did not respond to differences from) the prototypical vowel more often than the non-prototypical one. This result has been extended to adult subjects (Kuhl, 1991; Iverson & Kuhl, 1995).

What has subsequently become apparent is that the degree to which infants treat instances of a vowel distribution equivalently is critically dependent upon their experience with a particular language. Evidence supporting a role for learning by 6 months of age can be found in a study (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992) using the same paradigm as Grieser & Kuhl (1989) with infants from different language environments. According to this report, English infants displayed greater discrimination for changes between a "prototypical" version of the high front rounded Swedish vowel [y] and its variants than were obtained for differences between the "prototypical" English [i] and its variants. The complementary pattern was found for Swedish infants' responses.

Kuhl (1991) conceptualizes these discrimination data as a "perceptual-magnet effect" and suggests that infants come to internalize vowel category prototypes similar to those for adults, and that variants of the vowel category are perceptually assimilated to the prototype or "Natural Language Magnet" (Kuhl, 1993) to a greater degree than could be explained on psychophysical distance alone. Kuhl's (1993) "Native Language Magnet" (NLM) theory postulates that infants have innately-given natural auditory boundaries that partition the auditory space. These boundaries are taken to be species-general. With exposure to language, infants develop “natural magnets” that define their language’s phonetic categories.

These magnets "shrink" the auditory space in their vicinity allowing for the creation of equivalence classes for phonetic segments.

Thus, NLM suggests that the perceptual vowel space of the language perceiver is not solely a function of auditory processes but is also a specific function of language exposure. In particular, the space shrinks around representative exemplars. This complicates the predictions of phonetic inventories from perceiver characteristics as performed by Lindblom (1986), for example. Perceptual distinctiveness arises in great part, according to NLM, from processes of learning and not merely from acoustic difference. As space shrinks around prototypes, distinctiveness *within* a category decreases and distinctiveness *between* categories increases. Acoustic difference is less of a determiner of distinctiveness than suggested in models like Lindblom (1986).

It is probably unsurprising that exposure and learning have some effect on the notion of "perceptual distinctiveness" and NLM offers an interesting description of what this effect may look like. But is it the correct description of the influence of learning and exposure? One troubling aspect of this theory is the concept of the representative prototype. If the prototype is formed by representativeness then why do listeners place phonetic targets at extreme values in the formant space instead of at typical production values (Johnson, Flemming, & Wright, 1993)? In addition, NLM fails to predict typical vowel inventories. Presumably any perceptual space could result from appropriate input and distinctiveness would be maintained by the perceptual magnets. The modal phonetic systems with maximally-dispersed vowel types are not explicitly predicted by a perceptual-magnet approach. As we will see below, however, some learning models would predict dispersed vowel systems.

NLM has been quite influential. Demonstrations of "perceptual-magnet effects" have been described for several phonetic contrasts (e.g. Davis & Kuhl, 1994; Iverson, Diesch, Siebert, & Kuhl, 1994; Iverson & Kuhl, 1996) and "perceptual-magnet effects" have been modeled by various computational "neural network" approaches (Guenther & Gjaja, 1996; Lacerda, 1995). The following exposition offers some cautionary notes concerning NLM theory and the "perceptual-magnet effect". In particular, concerns will be presented about the stimulus set used in the experiments previously described; failed attempts at replicating some of the results will be reported; and potential problems with the methodology of "perceptual-magnet" effects will be discussed.

X.1.1 Stimuli Concerns

The first set of cautions about demonstrations of "perceptual-magnet effects" concern the stimulus sets. As noted previously, it is rather important that all of the stimuli in these experiments be perceived as members of the same phonetic class (e.g., [i]). However, the identity of the non-prototype [i] from Grieser & Kuhl (1989; Kuhl, 1991) has been called into question. In our lab, we have presented adult listeners with the Grieser & Kuhl [i] stimuli along the **P**-**NP** axis in a free identification task. Data pooled across 15 subjects are presented in Table I. Also presented are the identifications from one anomalous subject whose data were not included in the summary. The fifth stimulus corresponds to Grieser & Kuhl's "prototype", while the ninth stimulus is the "non-prototype". As can be seen, the non-prototype was identified as /i/ only 7.3 percent of the time. It appears to be a better /ε/ than an /i/.

Table I. Average percentage identifications for stimuli along F1-F2 axis containing **P** and **NP**. Stimulus 5 was the Grieser & Kuhl (1989) “prototype”. Stimulus 9 was the “non-prototype”. Data from an anomalous subject who found none of the stimuli to be compelling versions of /i/ are presented separately.

15 Subjects (pooled)

Stimulus Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
% /i/	95.3	96.7	96.7	96.7	98.0	90.7	70.7	23.3	7.3	1.3	0.0	3.3	0.7
% /ɪ/	1.3	0.0	0.7	0.0	1.3	2.0	8.7	18.7	18.0	14.7	16.0	12.7	6.7
% /ɛ/	0.7	0.0	0.7	0.0	0.7	2.7	9.3	25.3	39.3	46.7	47.3	43.3	35.3
% /e/	0.0	0.7	2.0	0.0	0.0	1.3	5.3	22.0	22.7	26.0	30.0	32.0	38.0
% /æ/	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	7.3	6.7	6.0	8.0	16.0
% /ʊ/	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.0	0.0	0.0	0.0	0.0
% None	2.7	2.7	0.0	3.3	0.0	3.3	6.0	6.0	5.3	4.7	0.7	0.7	3.3

Anomalous Subject

Stimulus Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
% /ɪ/	100	100	100	100	100	100	90.0	20.0	0.0	0.0	0.0	0.0	0.0
% /æ/	0.0	0.0	0.0	0.0	0.0	0.0	10.0	60.0	80.0	90.0	80.0	90.0	90.0
% None	0.0	0.0	0.0	0.0	0.0	0.0	0.0	20.0	20.0	10.0	20.0	10.0	10.0

This result has also been obtained by Sussman & Lauckner-Morano (1995). In their study, phonetically-trained subjects who were presented the entire Grieser & Kuhl [i] distribution labeled the "non-prototype" as /i/ only 8% of the time. Thus, it appears that infants in Grieser & Kuhl's "prototype" condition were performing intra-category discriminations ([i] versus [i]); whereas, infants in the "non-prototype" condition were performing *inter*-category discriminations ([i] versus not-[i]). In this light, the higher discrimination scores obtained for the "non-prototype" condition are unsurprising. Increased discriminability of cross-boundary pairs versus within-category pairs is one of the classical findings of speech perception (Pisoni, 1973).

Of additional concern is the typicality of the "prototype" stimulus. The duration of vowels in the Grieser & Kuhl (1989) study was 500 msec. This is clearly not typical of the duration of natural vowel utterances. Likewise, the monothongal "prototype" Swedish vowel /y/ used in the Kuhl *et al*. (1992) study was not a particularly compelling version of this typically dipthongized vowel. Because of constraints of stimulus construction, the frequency of F3 was too high to model natural productions of these vowels. While the "prototype" vowels may have been considered the most typical as compared to the rest of the set, they probably do not represent truly typical Swedish vowels. Given the importance of the prototype conception in Kuhl's NLM theory, this could present a problem. It appears that, in these experiments, the purported shrinking of the auditory space isn't around a prototype but around an exemplar which is fairly atypical. Given this, what is the ontological status of the “prototype”?

X.1.2 Replications

A second concern with data from the paradigmatic "perceptual-magnet" experiments arises from subsequent failures to replicate the adult goodness judgments presented in Kuhl (1991). The original data show that the "prototype" vowel was judged as the best member of the category [i], and the "goodness" judgments fall off symmetrically with distance from the "prototype". However, in a replication, Lively (1995) reported that, for his subjects, the “prototype” was not given the highest rating, and, in fact, the exemplar with the highest rating was the one with the highest F2 frequency. The ratings were clearly not symmetrical around the "prototype" vowel. Iverson & Kuhl (1995) and Lotto, Kluender, & Holt (1995) obtained similar results for "goodness" judgments on vowels from the **P-NP** diagonal. Vowels that were more extreme in F2 frequency were rated as the best members of the category [i].

These results are in agreement with those found by Johnson *et al*. (1993). Using a method of adjustment, they asked listeners to choose an acoustic target to match the vowel in a visually-presented word. The targets were more extreme in the F1-F2 space than were the produced vowels for these words. In particular, the target [i] had a higher F2 (and lower F1) than the produced (and presumably more representative) [i].

Once again, it is difficult to ascertain the theoretical importance of the "prototype" concept in these experiments, because the "prototype" vowel may not have been the best exemplar in the set. Later, we will return to these gradients in goodness judgments.

X.1.3 Methodology

There is an aspect of the methodology used in adult "perceptual-magnet" experiments which exacerbates concerns about stimulus identity presented above. In a typical experiment (e.g. Iverson & Kuhl, 1995; 1996) subjects are presented with stimuli in isolation and are asked to identify the stimulus and give it a "goodness" rating. Following this, subjects participate in a same-different discrimination task in which they are presented pairs of stimuli. Thus, the conditions for obtaining identification and goodness data (isolated stimuli) differ from those for obtaining discrimination scores (paired stimuli). This is a concern because there is ample evidence for context effects on identification of speech sounds. Whether it is called selective adaptation (Eimas & Corbit, 1973), vowel contrast (Johnson, 1990) or frequency (spectral) contrast (Lotto & Kluender, 1998), the perceived identity of a speech sound can be affected by contextual sounds (see Diehl, 1981 for review). It is possible that the perceived phonemic identity of exemplars in a discrimination experiment differ from their identity in isolation.

Given this concern, Lotto, Kluender & Holt (1998) conducted a modified replication of the experiments described in Iverson & Kuhl (1995). In the Iverson & Kuhl study, adult subjects were presented stimuli from the **P-NP** axis of the Grieser & Kuhl (1989) distributions. Subjects identified each of the vowels presented in isolation and provided a goodness rating. A second group of listeners participated in a discrimination task. In the "prototype" condition, the "prototype" vowel was discriminated from the six closest exemplars. In the other condition, the "non-prototype" vowel was discriminated from its six closest neighbors. The Iverson & Kuhl data revealed that the "prototype" was less discriminable from its neighbors than was the "non-prototype" from its neighbors -- the "perceptual-magnet effect". This difference in discriminability appeared to be greater than what one would predict from the identification functions of the isolated vowels.

In the Lotto *et al*. replication, the stimulus presentation for the identification/goodness tasks and for the discrimination task were identical. That is, subjects identified vowel stimuli in context of the same vowel that later would serve as the paired vowel in discrimination. In the "prototype" condition, vowels were identified in context of the "prototype" vowel. In the "non-prototype" condition, vowels were identified in context of the "non-prototype" vowel. The effects of context on identification of vowels presented in both conditions were rather staggering. In the presence of a relatively poor [i] exemplar ("non-prototype") these vowels were labeled as [i] 91% of the time. On the other hand these vowels were labeled as [i] only 26% of the time when paired with a relatively good [i] ("prototype"). In addition, goodness judgments were highly affected by context. Consistent with findings from a host of earlier studies (e.g. Thompson & Hollien, 1970), vowels were judged to be much better versions of [i] when they were in context of a rather poor version of [i]; actually, in context of a vowel that is not judged to be an [i] at all.

In the discrimination task, Lotto *et al*. obtained data very similar to Iverson & Kuhl (1995). The "prototype" condition, indeed, led to lower **d**' scores than did the "nonprototype" condition. Lotto *et al*. calculated the predicted generalization scores (percent misses) from the identification curves obtained in context conditions. These computations essentially amount to the probability that the two vowels would be labeled equivalently. This gives the predicted number of misses that would arise simply from differential identification (Liberman, Harris, Hoffman & Griffith, 1957). This doesn't mean that vowels are only discriminable insofar as they differ in phonemic label (Pisoni, 1973), but that one can predict differential effects arising solely because of labeling.

The predicted difference in generalization scores between "prototype" and "nonprototype" conditions was actually slightly larger than the empirically-obtained scores. That is, the empirically-derived higher discrimination for the "nonprototype" comparisons could be completely accounted for by labeling differences from changes in context. There is nothing left for NLM to explain. There is **no** "perceptual-magnet effect" here.

X.2 Alternatives to Perceptual Magnets

If NLM dissipates in the light of the preceding data, then are we left with a vapor lock when considering the role of learning in the structure of phonetic inventories? Perhaps not. The general ability of humans to learn probabilistic mappings between domains has been well documented and described in psychological literature (e.g. the statistical learning theory of Estes, 1950). The value of connectionist or neural network approaches which are epidemic in current psychological modeling is that they can duplicate much of the human ability to learn concepts based on statistical regularities of inputs (as opposed to learning by explicit rule). Acquiring a native phonetic inventory is a paradigmatic case of this kind of statistical learning. Varying acoustic forms are presumed to be mapped onto abstract forms, such as phonemes or morphemes, based on cues which are only probabilistically linked to the abstract form.

The lack of invariance in these mappings have often been taken as evidence that specialized mechanisms are required to correctly perceive speech and to acquire phonetic categories. For example, the extreme context sensitivity and variability of the acoustic realization of [d] was one of the main phenomena adduced to support the claims of a specialized "speech code" (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). However, it is just this kind of probabilistic multiple-attribute learning at which biological systems excel (Brunswick, 1955; Hebb, 1949). In fact, birds with quite small brains (compared to humans) can learn the category for [d] in spite of the inherent variability in acoustic characteristics (Kluender, Diehl & Killeen, 1987).

Thus, statistical learning theory and connectionist modeling may well inform us about the acquisition of phonetic indentification and the constraints that this acquisition places on the form of language in general and the membership of phonetic inventories in particular. There have been several attempts to use connectionist or statistical principles to model acquisition. Unfortunately these attempts have, for the most part, been based on obtaining patterns of responses that are coherent with patterns already derived by empirical tests. There rarely are *a priori* predictions of the form of human responses which are based on the constraints of the learning system itself. Some have even developed models of perceptual magnets (e.g. Guenther & Gjaja, 1996; Lacerda, 1995). In light of the data described earlier in this paper, these attempts may have been misled.

Besides falsifiable predictions based on general learning principles, we need sound methodologies for testing the validity of these predictions for speech stimuli. One possible methodology is the use of nonhuman animals in learning experiments with speech-sound distributions. Animal models are a natural choice considering that much of what we know about general learning systems has been gleaned from studies of animals in a variety of tasks. This historical connection to behaviorism leads to some dismissal of animal models of speech perception, but the benefits of animal research for questions of phonetic acquisition are several. One can ethically and practically control the exact form and amount of experience that animals have with speech-sound distributions. In addition, one can sample performance during acquisition much more finely than is reasonable with human infants.

Kuhl (1991) demonstrated that monkeys that had *no previous experience with phonetic distributions* showed no perceptual-magnet effect for vowel stimuli. It is important to note that this result demonstrates that previous findings with human adults and infants are not explainable by "basic auditory processes". It *does not*, in itself, support a notion that the acquisition of linguistically-appropriate behavior is due to species-specific processes.

In contrast to Kuhl's experiment, Lotto *et al*. (1995) studied the patterns of responses to vowel sounds by an animal after substantial exposure to distributions of these vowels. Starling were trained to peck a key when presented a vowel from one distribution (e.g. [i]) and to refrain from pecking the key when presented with vowels from a competing distribution (e.g. [ɪ]). After a mere 101 hours exposure to these distributions, the birds readily generalized their behavior to novel tokens drawn from the distributions. One of the more salient aspects of the birds' patterns of response was that vowels that were furthest from the competing distribution were responded to most heavily. For example, a bird trained to peck to sounds from an [i] distribution would peck most strongly to vowels with high F2 frequencies (away from the [ɪ] distribution). This result may be predicted from classic findings in the general learning literature. It has been well-established that the response to a positive stimulus is often skewed away from a negative stimulus in discrimination learning (e.g. Hansen, 1959).

This finding is particularly interesting in light of the data on human goodness judgments of vowel distributions. As noted above, human adults also find tokens with extreme F2 values to be better exemplars of [i] than more representative values (Lively, 1995; Lotto *et al*. 1995). It is quite possible that this contrastive pattern in human responses is due to the structure imposed by general learning processes. As one is exposed to speech-sound distributions, the tokens which are considered representative become further removed from each other in acoustic space because of the constraints of learning. In a sense, a more discriminative vowel space may be the result of learning and not simply to the virtues of maximizing auditory distinctiveness as proposed by such theories as "Auditory Dispersion" (Lindblom, 1986). In addition, the findings from the bird study offer a potential explanation of the fact that "phonetic targets" for production are more extreme than typical acoustic realizations of phonetic segments (Johnson *et al*., 1993). Lindblom's (1990) 'Hyper- and Hypoarticulation' theory proposes that the targets are maximally distinctive because they are the realizations of the most effortful, hyperarticulated speech. We may speculate that these targets may be maximally distinctive because general learning processes result in a disperse pattern of responses ('categories') as they relate to competing responses.

At this point these suggestions are highly speculative and the dynamics of these purported processes are underspecified. A combination of animal models, human adult and infant response data to speech and nonspeech and computational models will be needed to make more particular predictions about the structure imposed by general processes of statistical learning. The resulting increased precision may lead to models which can specify phonetic inventories (and structure at other linguistic levels) based on constraints imposed by the peripheral auditory system, articulatory production *and* the processes of perceptual learning.

References

Brunswick, E. 1955. Representative design and probabilistic functionalism: A reply. *Psychological Review* 62: 236-242.

Davis, K. & P. K. Kuhl. 1994. Tests of the perceptual magnet effect for American English /k/ and /g/. *Journal of the Acoustical Society of America* 95: 2976.

Diehl, R. L. 1981. Feature detectors for speech: A critical appraisal. *Psychological Bulletin* 89: 1-18.

Diehl, R. L. & K. R. Kluender. 1989. On the objects of speech perception. *Ecological Psychology* 1: 121-144.

Eimas, P. D. & J. D. Corbit. 1973. Selective adaptation of linguistic feature detectors. *Perception & Psychophysics* 13: 247-252.

Estes, W. K. 1950. Toward a statistical theory of learning. *Psychological Review* 57: 94-107.

Grieser, D. & P. K. Kuhl. 1989. Categorization of speech by infants: Support for speech-sound prototypes. *Developmental Psychology* 25: 577-588.

Guenther, F. H. & M. N. Gjaja. 1996. The perceptual magnet effect as an emergent feature of neural map formation. *Journal of the Acoustical Society of America* 99: 2590.

Hansen, H. M. 1959. Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology* 58: 321-372.

Hebb, D. 1949. *Organization of Behaviour*. New York: Wiley.

Iverson, P., E. Diesch, C. Siebert & P. K. Kuhl. 1994. Cross-language tests of the perceptual magnet effect for /r/ and /l/. *Journal of the Acoustical Society of America* 96: 3228.

Iverson, P. & P. K. Kuhl. 1995. Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *Journal of the Acoustical Society of America* 97: 553-562.

Iverson, P. & Kuhl. 1996. Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. *Journal of the Acoustical Society of America* 99: 1130-1140.

Johnson, K. 1990. Contrast and normalization in vowel perception. *Journal of Phonetics* 18: 229-254.

Johnson, K., E. Flemming & R. Wright. 1993. The hyperspace effect: Phonetic targets are hyperarticulated. *Language* 69: 505-528.

Kluender, K. R., R. L. Diehl & P. R. Killeen. 1987. Japanese quail can learn phonetic categories. *Science* 237: 1195-1197.

Kuhl, P. K. 1983. Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development* 6: 263-285.

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

Kuhl, P. K. 1993. Innate predispositions and the effects of experience in speech perception: The native language magnet theory. In B. de Boysson-Bardies (ed.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*. The Hague: Kluwer Academic Publishers, 259-74.

Kuhl, P. K., K. A. Williams, F. Lacerda, K. N. Stevens & B. Lindblom. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255: 606-608.

Lacerda, F. 1995. The perceptual-magnet effect: An emergent consequence of exemplar-based phonetic memory. In K. Elenius & P. Branderud, (eds.), *Proceedings of the XIIIth International Congress of Phonetic Sciences*, *Vol. 2*. Stockholm: KTH and Stockholm University, 140-47.

Liberman, A. M., F. S. Cooper, D. P. Shankweiler & M. Studdert-Kennedy. 1967. Perception of the speech code. *Psychological Review* 74: 431-461.

Liberman, A. M., K. S. Harris, H. S. Hoffman & B. C. Griffith. 1957. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 54: 358-368.

Liljencrants, J. & B. Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contast. *Language* 48: 839-862.

Lindblom, B. 1986. Phonetic universals in vowel systems. In J. J. Ohala & J. J. Jaeger (eds.) *Experimental Phonology*. Orlando, Florida: Academic Press, 13-44.

Lindblom, B. 1990. Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (eds.) *Speech Production of Speech Modeling*. Dodrecht: Kluwer, 403-39.

Lindblom, B. & I. Maddieson. 1988. Phonetic universals in consonant systems. In L. M. Hyman & C. N. Li (eds.) *Language, Speech and Mind*. New York: Routledge, 62-78.

Lively, S. E. 1993. An examination of the perceptual magnet effect. *Journal of the Acoustical Society of America* 93: 2423.

Lotto, A. J. & K. R. Kluender. 1998. General contrast effects in speech perception: Effect of preceding liquid on stop consonant identification. *Perception & Psychophysics* 60: 602-619.

Lotto, A. J., K. R. Kluender & L. L. Holt. 1995. Animal and computational models of development of graded vowel categories. *Journal of the Acoustical Society of America* 98: 2965.

Lotto, A. J., K. R. Kluender & L. L. Holt. Forthcoming. Depolarizing the perceptual magnet effect. *Journal of the Acoustical Society of America*.

Ohala, J. J. 1974. Experimental historical phonology. In J. M. Anderson & C. Jones (eds.) *Historical Linguistics II: Theory and Description in Phonology*. Amsterdam: North Holland Publishing Company, 353-89.

Peterson, G. E. & H. L. Barney. 1952. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24: 175-184.

Pisoni, D. B. 1973. Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics* 13: 253-260.

Sussman, J. E. & V. J. Lauckner-Morano. 1995. Further tests of the perceptual magnet effect in the perception of [i]: Identification and change/no-change discrimination. *Journal of the Acoustical Society of America* 97: 539-552.

Stevens, K. N. 1989. On the quantal nature of speech. *Journal of Phonetics* 17: 3-45.

Thompson, C. L. & H. Hollien. 1970. Some contextual effects on the perception of synthetic vowels. *Language and Speech* 13: 1-13.