Incorporating principles of general learning in theories of language acquisition

Lori L. Holt[1], Andrew J. Lotto[2], & Keith R. Kluender[1]
[1]University of Wisconsin-Madison
[2]Loyola University-Chicago & Parmly Hearing Institute

Introduction

Before they become native speakers, infants are native listeners. A great deal of evidence supports the observation that, during their first year, human infants come to respond to elements of language in a manner appropriate to the language environment in which they are being reared. Infants' native-language-appropriate perception extends across phonetic (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker & Tees, 1984a), prosodic (Hirsh-Pasek, Kemler Nelson, Jusczyk, Wright Cassidy, Druss, & Kennedy, 1987; Jusczyk, Cutler, & Redanz, 1993), and phonotactic (Friederici & Wessels, 1993; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993) levels, strongly suggesting that months before they utter their first words, infants have begun their initiation into a language community.

These findings offer glimpses of the developmental unfolding of speech-perception abilities and generate many exciting, central questions as to the processes involved in language acquisition. Among these questions, perhaps none creates more theoretical fury than the controversy between "nature" and "nurture". The attempt to resolve relative contributions of innate predisposition and experience-driven organization of the initial state is paramount in theoretical importance for language acquisition. The rapidity and seeming ease with which infants come to perceive elements of language in a manner appropriate to their native tongue has led some theorists to propose that innate language-acquisition mechanisms guide development.

Obviously, even staunch nativists acknowledge that language experience is pivotal. Language skills cannot unfold entirely according to a strict blueprint of human maturation because infants, after all, become native speakers and listeners of the language indigenous to the linguistic community in which they are raised. Experience, at the very least, shapes language acquisition. However, the position of experience and learning-based accounts of language acquisition has often been secondary to presumed innate mechanisms in theoretical accounts of language acquisition. Eimas (1987), for example, suggests that infants would find the sort of proficiency they demonstrate in their native-language-appropriate perception "inordinately difficult to learn, to say the least" (pg. 228). Typically, theories of language acquisition have concluded that "simple" mechanisms such as learning are insufficient to account for the wealth of linguistic complexity involved in learning a language. Such views may arise from downward generalization of classic arguments concerning syntax (e.g., Chomsky, 1959; Skinner, 1957).

However, with respect to the developmental course of language-appropriate perception, theorists may have underestimated the potential power of learning mechanisms coupled with a tremendously statistically-rich input like language. Learning may play a much broader role in language development than has been postulated.

Few theories or empirical investigations have examined the degree to which elements of language acquisition may be accounted for by general principles of learning without reference to innate constructs or species-specific mechanisms. This is undoubtedly due, at least in part, to methodological difficulties. Ethical considerations forbid direct manipulation of infants' language experience and experiments with adult subjects are not necessarily relevant because adults are already polished native listeners.

Fortunately, experience-based hypotheses need not be abandoned for lack of an appropriate empirical test. Animal models provide an excellent means by which to explore the provocative and testable predictions of an experience-driven account of language acquisition. A good deal of evidence suggests that essential characteristics of human speech perception are mirrored by nonhuman listeners. For example, nonhuman animals have been shown to respond to speech categorically (Kuhl & Miller, 1975; 1978; Morse & Snowdon, 1975; Waters & Wilson, 1976), to exhibit speech context effects (Dent, Brittan-Powell, Dooling, & Pierce, 1997; Lotto, Kluender, & Holt, 1997), to be sensitive to acoustic trading relations (Kluender & Lotto, 1994), and to make use of multiple acoustic attributes in classifying speech (Kluender, Diehl, & Killeen, 1987). Nonhuman animals are unlikely recipients of innate predisposition to acquire language or of specialized neural machinery to perceive sounds produced by a human vocal tract. Moreover, one would expect that general learning mechanisms possessed by nonhuman animals also are available to human infants. With these qualities, nonhuman animal models offer a well-suited, ethically-sound population with which to probe perceptual ramifications of experience with speech. Evidence that a nonhuman species exhibits perceptual organization as a result of experience with language elements provides support for the claim that language-appropriate perception can arise from learning without innate language-acquisition mechanisms and a means by which characteristics of learning mechanisms can be revealed.

The approach advocated here is to determine to what degree infants' native-language appropriate perceptual capacities can be accounted for by general learning mechanisms before postulating computationally-complex, language-specific, or species-specific alternatives. Nonhuman animal models provide the most fertile experimental ground for examining these questions. Here, several grounding assumptions of this approach are made explicit; preliminary evidence demonstrating the efficacy of animal models in investigating effects of experience on one element of language acquisition (phonetic equivalence-class acquisition) is presented; predictions of a learning-based account for another element of phonetic development are outlined; and, a means to test these predictions is suggested. The main emphasis here is upon discovering the processes that underlie infants' developing language-specificity of *phonetic* perception, but the ideas presented rely only upon statistical regularities in input and learning coupled with general auditory capacities. As such, these hypotheses also are applicable, in principle, to other arenas of language-appropriate perceptual development.

Assumptions: Statistical Regularity and General Learning

Fundamentally, assumptions of the present approach to investigating infants' emerging native-language specificity in speech perception are grounded in statistical learning (in the spirit of Estes, 1950) whereby the burden of explanation lies not in hypothetical innate constructs but in environmental regularities and organisms' ability to make use of them via learning. An account of this sort is feasible because biological organisms are extraordinarily successful at recovering statistical patterns of their environment. Successful perceptual systems, in general, come to complement regularities present in the environment (Young, 1962).

Language is a particularly fecund source of statistical regularities. Across languages, rhythm patterns, phonetic inventories, prosodic melodies, and phonotactic orderings exhibit inherent regularities. Phonetic inventories provide a well-established example. The phonetic patterns infants must discover in acquiring language differ greatly across native languages. Fortunately, there is substantial within-language regularity in phonetic classes to guide infants in phonetic acquisition. This point is illustrated by an early inventory of cross-language stop consonant voice-onset-time (VOT) productions. In measuring acoustic indications of VOT, Lisker and Abramson (1964) observed typical between- and within-speaker variation of VOT values. However, along with this variability, they also reported very regular underlying acoustic patterns for the phonemes of each language. Within a language, estimated VOT values for productions of a particular phoneme tended to cluster around a mean VOT value that occurred most frequently across productions. There was, in addition, variance in VOT such that values adjacent to the mean value were also observed, but less frequently. Stated another way, the distribution of estimated VOT values for a particular phoneme approximated a normal distribution.

Precise characteristics of the acoustic distributions are relatively unimportant. What is critical is the observation that observed phonetic acoustics of a language, although variable, are comprised of statistical regularities -- systematic distributions of acoustic patterns corresponding to the phonetic segments a language employs. With this information available, infants can learn the probabilistic mappings from auditory form to functional elements such as phonemes or syllables. Acoustic-phonemic mappings are probabilistic in nature. They do not provide a unique invariant acoustic pattern corresponding to a particular phoneme. However, they do provide just what an infant language-learner requires -- reliable statistical information as to the phonetic patterning of the native language. The claim that acoustic structure of phonemes, although stochastic in realization, has underlying structure within a language is perhaps unsurprising in that speech, as a communication system, depends upon the distinctiveness of its constituent sounds to convey information. However, it is precisely this point that provides the means by which infants might come to exhibit language-appropriate perception through general learning processes sensitive to statistical patterns of language input. The patterns underlying phonetics, and other elements of language, create a statistically-robust substrate upon which learning processes may operate.

If infants can make use of these statistical regularities in discovering the organization of their particular native language, then a learning-based account of the developmental course of native-language-appropriate perception may be tenable. There is intriguing evidence that infants may indeed be quite sensitive to statistical regularities of sound. Quite early, infants recognize relative frequencies of phonotactic patterns of their native language. For example, when presented with phonotactically-permissible phonetic strings, nine-month-old infants prefer to listen to patterns that are used more frequently in their native language than they do to more infrequent strings (Jusczyk, Luce, & Charles-Luce, 1994). Infants also appear to be sensitive to patterns imbedded within a continuous stream of sounds. Eight-month-old infants can segment pseudo-words from a continuous stream of nonsense speech even when the only cues to word segmentation are statistical regularities in relative occurrence of syllabic elements (Saffran, Aslin, & Newport, 1996). Given the absence of other segmentation cues in this study, infants appear to have been reliant on statistical learning of transitional probabilities between syllables. Most impressively, these infants learned statistical distributions of syllable occurrence and successfully applied it to word segmentation after a mere two minutes of exposure (Saffran et al., 1996)! In an extension of these results, infants were shown also to learn readily statistical correlation of nonlinguistic (tonal) sound patterns (Johnson, Saffran, Aslin, & Newport, 1998). Taken together, these data suggest that infants are adept at statistical learning of language elements, or other statistically-structured input.

These results are promising for learning accounts of development of native-language-appropriate speech perception. Infants appear to be sensitive to the rich internal structure of their native language and, indeed, to more impoverished laboratory inventions of patterned input. From a very early age, infants posses the ability to recognize, learn, and respond to regularities in input. These are precisely the skills that would encourage native-language-appropriate perceptual organization via learning.

It is important to note that, at this point, no specific learning theory is implicated by the general hypothesis that learning plays a substantial role in development of native-language-appropriate perception. It is perhaps accurate to refer to the type of learning implicated as *perceptual learning* in that, by the account advocated here, infants' changes in perception are attributed to learning as a consequence of experience or, alternatively, as *statistical learning* in that infants' perception comes to complement statistical regularities of their native language. Importantly, putative learning is not necessarily dependent upon classical associative models in which a stimulus is paired with reinforcement. Rather, learning may arise from subtler self-organization of perception which, in principle, requires no reinforcement to occur (e.g., Kohonen, 1987).

Preliminary Evidence: Phonetic Equivalence-Class Development

Whereas data demonstrating infants' excellent abilities to recognize and make use of statistical regularities in input are greatly encouraging for models that suggest a broader role for general learning processes in language acquisition, it is difficult to unravel potential mechanisms with studies involving human infant subjects. Experimental constraints introduced by infants' napping, fussiness, and short attention spans significantly limit the number of data points that may be collected from a single individual. Coupled with difficulties involved in longitudinal infant studies, this factor has contributed to the unavailability of a large longitudinal corpus of infant data. Rather than providing a fluid, finely sampled "motion-picture" of the progress of unfolding infant speech perception, the presently-available infant data offer information more similar to sporadic low-resolution snapshots across time. Animal models, where control of experience is entirely in the hands of the experimenter and many data points can be collected across time, provide a means by which to begin to resolve whether learning mechanisms could play a significant role in these infant behaviors and, if so, to understand the characteristics of these mechanisms.

Using an animal paradigm in this way, Kluender, Lotto, Holt, and Bloedel (1998a) have modeled one aspect of infants' emerging language-appropriate perception: phonetic equivalence-class acquisition. By six months of age, infants respond to acoustically-different vowel sounds in a manner that respects their functional equivalence within the native language (Kuhl, 1983; Kuhl et al., 1992). Kuhl et al. (1992), for example, trained infants to turn their heads toward a loudspeaker in response to a vowel change from a very good ("prototypical") version of a vowel to a less-exemplary instance of the same vowel. For novel native and nonnative vowel pairs that were equated in acoustic and auditory change, six-month-olds were more likely to respond to changes across nonnative vowel pairs than to changes in native vowel pairs. That is, infants were more likely to treat as functionally-equivalent native vowels with which they had had most experience. Dissimilarly, Kuhl (1991) reported that rhesus monkey subjects, lacking experience with speech, fail to respond to changes in vowels in a manner that suggests functional equivalence within a class. Thus, it appears that primate mammalian systems do not, in and of themselves, give rise to functional vowel equivalence classes. As Kuhl (1991) noted, these data suggest either that human infants are biologically endowed with species-specific mechanisms that guide development of phonetic equivalence classes or that the effects witnessed in infants are due to experience listening to a native language.

Use of nonhuman animals allows an explicit test of this dichotomy. To this end, Kluender et al. (1998a) used the avian species European (Common) starling (*Sturnus vulgaris*) in a vowel-learning experiment. Birds were trained to discriminate vowel tokens drawn from stylized distributions either of English vowels /i/ and /ɪ/ or of Swedish vowels /y/ and /ʉ/. Following training with a subset of either English or Swedish vowel exemplars, avian subjects were tested on novel vowels. For these novel test tokens, "English" starlings failed to respond to differences across test tokens (belonging to the same functional distribution) that were responded to differentially by "Swedish" birds (for whom the stimuli belonged to different functional distributions). Likewise, each "Swedish" bird failed to respond to differences in test tokens that "English" birds responded to differentially. Thus, experience with a subset of either "English" or "Swedish" training stimuli influenced birds' responses to novel test pairs. These avian subjects appear to have learned the phonetic equivalence of stimuli drawn from a distribution. Notably, starlings' performance was based on only about 100 hours of experience.

What is particularly interesting is that the functional equivalence-classes of the birds, like those of human adults and infants, were not homogeneous. Birds did not treat all novel vowel exemplars as equivalent. Rather, birds tended to respond most to vowels that were maximally distinct from the opposite "native" functional equivalence class. This response gradient is typical of human phonetic perception (as evidenced in "goodness" judgements) and can be predicted from very early principles of general learning (e.g., Spence, 1936; 1937; 1952; 1960) whereby the response pattern to an excitatory distribution is skewed away from an inhibitory stimulus class.

On nearly every account, there was striking agreement between measures of starling peck rates and human goodness judgments of the same vowels. This avian experiment suggests that it is possible to capture the process of infants' developing perceptual organization for speech sounds with little more than fundamental principles of learning. As such, these data are encouraging for animal models of human infant language acquisition. They also suggest that it is feasible to model human infant performance by exposing nonhuman subjects to statistical distributions of speech sounds.

A Related Question: Nonnative Speech Discrimination

The results of Kluender et al. (1998a) suggest that general learning principles may be able to account for substantial amounts of the structure of infants' development of native phonetic equivalence classes. Learning may prove to be fundamental in other arenas of language acquisition as well; studies that continue investigation of the role of learning in language acquisition are needed. The statistical-learning approach advocated here may prove to be a useful paradigm to guide further investigations because it suggests thoroughly testable predictions. In support of this claim and continuing with phonetic perception as a model case, a classic demonstration of the influence of native language experience in infants' perception is outlined and discussed.

Werker and her colleagues (Werker, Gilbert, Humphrey, & Tees, 1981; Werker & Logan, 1985; Werker & Lalonde, 1988; Werker & Tees, 1984a; 1984b) have contributed substantial evidence that, as a function of language experience, infants' tendency to respond to differences between nonnative speech sounds attenuates. In a series of studies, Werker and colleagues have exploited the fact that speakers of English and Hindi use place-of-articulation somewhat differently for stop consonants. In English, three places of articulation are used for voiced stop consonants: labial, alveolar, and velar ([b], [d], and [g] respectively). In Hindi, four places of articulation are used: labial, dental, retroflex, and velar ([b], [d5], [d8], and [g] respectively). Werker and Lalonde (1988) created a single synthetic series that varied perceptually from [b] to [d] (for native English adults) and from [b] to [d5] to [d8] (for native Hindi adults). Using a reinforced head-turn paradigm, they found that 6- to 8-month-old infants from English-speaking families responded to changes in stimulus tokens that perceptually crossed the English phonetic boundary between [b] and [d] and also responded to changes across the Hindi phonetic boundary separating [d5] and [d8]. A separate group of 11- to 13-month-old infants from English-speaking homes responded reliably to the English [b]-[d] contrast, but not to the Hindi [d5]-[d8] distinction. In addition, adult English speakers also demonstrated difficulties in discriminating the Hindi contrast. These results appear very robust; analogous evidence has been found for other cross-language distinctions (Best, McRoberts, LaFleur, & Silver-Isenstadt, 1995; Polka & Werker, 1994; Trehub, 1976; Werker & Tees, 1984a). Overall, 6- to 8-month-old infants more reliably respond to nonnative distinctions than do 11- to 13-month-olds or adults, suggesting that as infants experience speech sounds appropriate to their native language, they become unlikely or unable to discriminate nonnative phonemes.

Processes guiding these developmental changes remain unclear. However, language exposure, in itself, appears to be insufficient to account for the emergence of language-appropriate perception. Two pieces of evidence support this conclusion. Best, McRoberts, and Sithole (1988) have demonstrated that monolingual English adults and infants up to 14 months (the oldest infants tested) reliably discriminate click consonant contrasts from the Zulu language.[1] Clicks fall entirely outside typical English-language experience, yet English-learning infants and even English adults are able to discriminate click consonants. A second observation is that, in some cases, infant listeners *do* experience nonnative sounds as allophonic variants of native phonological classes (e.g., MacKain, 1982). In the case of place-of-articulation, for example, English infants are likely to experience instances of the English phonological class [d] that, in their realizations, very closely approximate good examples of the Hindi consonant

[d5] (e.g., "draw"). Yet, even with this exposure, English infants' ability to discriminate Hindi consonants changes with development.

If mere exposure is insufficient to account for these data, what is the nature of processes underlying this developmental change? Learning-based accounts respectful of the abundant statistical regularities available in language may provide a means by which to begin to resolve this question. For a statistical-learning account of infants' declining sensitivity to nonnative speech contrasts, the pivotal inquiry is whether experience with distributions of speech sounds is sufficient to encourage observed trajectories of nonnative speech discrimination. Like any adequate model, a learning-based account must be able to capture observed infant behavior. Minimally, models must account for the following observations:

(1) Infants with relatively less native-language experience discriminate nonnative contrasts.
(2) With more native-language experience, infants become unlikely or unable to discriminate nonnative contrasts.
(3) Those contrasts where discrimination maintains conform to native-language phonemic classes.
(4) Mere exposure does not appear to be adequate. Infants may continue to discriminate nonnative contrasts that are acoustically dissimilar to native phonemic classes and infants may fail to discriminate nonnative contrasts to which they have been exposed.

Relying upon principles of general learning and auditory capacities, the approach offered here provides testable predictions concerning each of these central observations.

In addressing young infants excellent nonnative-contrast discrimination, a general learning and auditory model suggests that early in infancy, when infants have had less native-language experience, speech discrimination is accomplished primarily by general auditory mechanisms, without much influence of learning. Performance arising from pure auditory capacities would, in general, predict excellent nonnative speech discrimination regardless of linguistic experience because languages tend to choose highly-discriminable speech contrasts for their phonetic inventories (e.g., Diehl & Kluender, 1989; Diehl, 1991). Furthermore, there is ample evidence that nonhuman animals, lacking experience with speech, readily discriminate speech contrasts (see Kluender, Lotto, & Holt, 1998b). An auditory-based account of early capacities is thus in line with the observation that younger infants effortlessly discriminate both native and nonnative speech contrasts.

With increased experience, infants meet a developmental milestone after which they are unable or less likely to discriminate nonnative speech contrasts. Moreover, the sensitivities that infants maintain correspond to native phonemic contrasts. The present approach suggests that this developmental change may be accounted for by infants' development of functional equivalence classes characteristic of their native language. Kluender et al. (1998a) demonstrated with an avian model that phonetic equivalence classes might be formed by very general learning mechanisms. If infants' experience with statistical patterns of a native phonetic inventory likewise is sufficient to form functional equivalence classes, then this perceptual organization may have implications for nonnative speech discrimination.

As infants acquire functional equivalence classes of native speech sounds, they begin to ignore acoustic differences in deference to functional similarity (e.g., Kuhl, 1983; Kuhl et al., 1992). Failure to discriminate nonnative contrasts that are acoustically similar to native phonemes may be brought about by infants' emerging inclination to categorize speech.[2] For example, consider the behavior of English-learning infants examined by Werker and Lalonde (1988). Acoustically, the Hindi consonants [d5] and [d8] are fairly similar to the English consonant [d]. If older English-learning infants have begun to develop a functional equivalence class corresponding to [d], then they are more likely than their 6-month-old counterparts to categorize examples of [d5] and [d8] as variants of the [d] equivalence-class. In categorization, discriminably-different examples are treated as functionally similar. It is possible that infants' failure to discriminate nonnative speech contrasts is a consequence of learning to equate speech exemplars according to native functional equivalence classes arising from experience with statistical regularities of phonetic experience. The ability to discriminate nonnative speech contrasts may not be so much *lost* as it is *preempted* by development of functional equivalence classes.

This hypothesis is in line with the additional observation that adults, when tested in different tasks, appear to be able to apply different processing strategies in discriminating nonnative speech contrasts. With appropriately sensitive tests, it is sometimes possible for adults to demonstrate discrimination of nonnative speech sounds (Werker & Tees, 1984a; Werker & Logan, 1985). In addition, this hypothesis may account for seemingly contradictory evidence that infants and adults maintain the ability to discriminate nonnative sounds very dissimilar from their native phonetic experience (Best et al., 1988). If a nonnative contrast is adequately distinct so as not to be sufficiently acoustically-similar to any native functional equivalence-class to be satisfactorily categorized, then the present model would suggest that auditory mechanisms continue to dominate perception.[3] Acoustically, the Zulu clicks employed by Best et al. (1988) likely were sufficiently acoustically distinct from any English phoneme to preclude generalization to a native functional equivalence class.

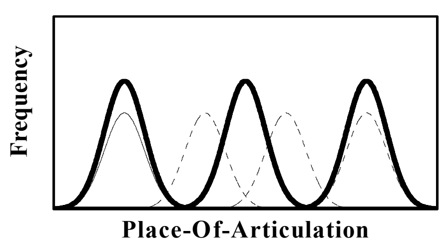
In sum, a statistical-learning account suggests that young infants exemplary nonnative contrast discrimination is a consequence of reliance upon primarily auditory mechanisms. With increasing experience with their native language, infants are hypothesized to develop functional equivalence classes for native phonemes. The mechanisms involved in this perceptual organization are postulated to be based on relatively general learning mechanisms sensitive to statistical distributions underlying native phonology. It is contingent upon development of these classes that infants become less likely to respond to nonnative contrasts. The ability to discriminate nonnative sounds may not be lost *per se*, but discrimination becomes more unlikely as responses are driven predominantly by function.

An Experimental Methodology

To determine whether these hypotheses withstand empirical scrutiny, a paradigm in which to test them is needed. Animal models offer an excellent means to accomplish this task. If experience with distributions of speech sounds is sufficient to encourage developmental changes like those observed in infants, then nonhuman animal speech discrimination patterns should be predictable from characteristics of the input distributions they experience.

In experimentally modeling these distributions, it is reasonable to assume that, statistically, experience with instances of a phoneme approximates a normal distribution. This assumption is supported, for VOT contrasts, by the acoustic measurements of Lisker and Abramson (1964). Thus, in absence of rigorous cross-language production measurements across many speakers and multiple contrasts, assuming approximate normality appears defensible. Following this assumption, phonetic experience of place-of-articulation for English-learning infants may be approximated by the dark, thick lines illustrated in Figure 1. Here, the abscissa represents the place-of-articulation dimension. Experimentally, this dimension may be approximated with a finely-sampled series of synthetic stop consonants varying perceptually for English adults from [ba] to [da] to [ga]. The ordinate represents relative frequency-of-occurrence of each of these stimuli. Assuming normality, these three distributions model relative occurrence of acoustically-distinct instances of the English phonemes [ba], [da], and [ga]. As discussed in reference to acoustic analyses of VOT (Lisker & Abramson, 1964), each distribution is assumed to represent relative frequency-of-occurrence of multiple productions of an English phoneme along place-of-articulation. English parcels place-of-articulation into three distinct classes and so there are three modes.

Figure 1

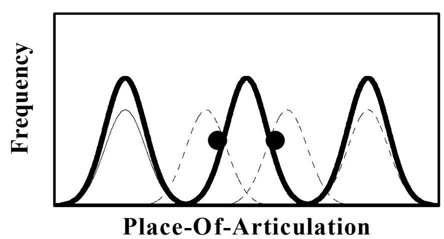


Likewise, the lighter, thinner lines of Figure 1 illustrate theoretical statistical distributions for Hindi phonological contrasts along the place-of-articulation dimension (i.e., [ba], [d5a], [d8a], and [ga]). Note that relative frequencies of stimuli in Hindi distributions are lower than those in English distributions. This is to assure that animals experiencing either set of distributions receive equivalent amounts of speech exposure; total area under each set of distributions is equal. Random sampling of either English or Hindi distributions in stimulus presentation provides a patterned input to animal subjects that approximates experience of English and Hindi infants. In addition, in experiments of this design, a control group of animals may experience stimuli sampled with equal probability along the place-of-articulation dimension so that animals receive no statistical patterning of input. With this very precise control over experienced statistical patterns of input across conditions, it is possible to test learning-based predictions that are simply unrealizable with infant subjects.

According to a learning-based account, animals' initial discrimination of speech, prior to substantial experience, should be predicted best by auditory/acoustic distance and any auditory discontinuities that may exist (e.g., Dooling, Best, & Brown, 1995; Kuhl & Miller, 1975; 1978; Kuhl & Padden, 1983). With increasing experience with distributions of speech sounds modeling a native-language environment, functional equivalence classes should be formed in a manner predictable from statistics of input. For animals experiencing stimuli drawn from distributions modeled after English, three functional equivalence classes should emerge; animals assigned to experience stimuli drawn from Hindi distributions should develop four functional classes; and, lacking structured input, control animals should fail to display functional equivalence. This prediction arises from the very simple hypothesis that, given a distribution of stimuli, organisms tend to parcel input in a manner that accounts for the most variance. This is precisely what self-organizing systems do (e.g., Kohonen, 1987)

If predictions of a general learning and auditory account are correct, corresponding with this, "English"-learning animals should fail to respond to differences in some critical stimuli to which "Hindi"-learning animals continue to respond. For example, consider the task of discriminating two stimuli that have the dual characteristics lying *within* an English distribution and also lying *across* two of the Hindi distributions. In the distributions illustrated in Figure 1, there is a stimulus combination that fits this description and also has the characteristic of being equivalent in frequency-of-occurrence for "Hindi" and "English" experimental populations. This pair is shown in Figure 2 where the stimuli are represented as dots on the distributions. For experienced "English"-learning animals, it is predicted that animals will not respond to a change between these stimuli because, as members of the same English distribution, the stimuli have come to be treated as functionally equivalent. Conversely, it is predicted that "Hindi"-learning animals *will* respond to this change because the stimuli cross Hindi distributions and are thus not treated as functionally equivalent. This example presents an important test of the role of mere exposure in language-appropriate development. Although "English"- and "Hindi"-learning animals have been exposed to these stimuli equally, the statistical structure of their global experience encourages very different responses whereas exposure predicts no difference between "Hindi" and "English" animals' responses because experience is equivalent.

Figure 2



With this sort of animal model, it will be possible to directly test each of the predictions outlined in the previous section. Furthermore, with this methodology, it will be feasible to examine in greater depth the effects of changes in statistical characteristics of input upon behavior and the developmental course of involved processes.

Conclusions

Determining the relative contribution of learning mechanisms to infants' development of native-language appropriate perception is fundamental to any theoretical model of language acquisition. Here, it has been argued that learning may play a much broader role in language acquisition than is typically suggested. The approach recommended is to reserve assertions of innate language-acquisition mechanisms until the explanatory potential of more general learning mechanisms has been exhausted. Scientific demands of parsimony and generality demand that learning, as a mechanism of developing language-appropriate perception, be given due consideration before postulating more elaborate species-specific and linguistically-designated mechanisms.

There is mounting evidence that infants, in fact, are capable of learning fairly subtle statistical patternings of language elements (e.g., Juscyzk et al., 1994 Saffran et al., 1996). There remains a great deal of future investigation necessary to clearly elucidate just how much of language-appropriate perception may arise solely from learning mechanisms, but infants' demonstrated sensitivity to statistical patterns, paired with the richness of linguistic experience, suggests that learning may play a substantial role. Although the present arguments have been directed primarily toward development of phonetic equivalence-classes, it is significant to note that this approach has important theoretical and empirical implications for other elements of language acquisition as well. For example, learning-based theories that embrace the information available in statistical regularities of language have promise in investigations of phonotactic or prosodic development. Across many elements of language, statistical regularities provide precisely the sort of multiply-correlated attributes that biological systems learn best (e.g., Hebb, 1949). Experience with this richly-complex input over the first months of life may substantially structure infants' language-appropriate perceptual capacities. However, this remains largely an empirical question. Fortunately, animal models are proving to serve as good models of infants' development (e.g., Kluender et al., 1998a; Kluender et al., 1998b).

In addition, the present approach promises to make contributions to computational modeling of infant language acquisition. The rather coarse sampling of available infant data has so far precluded attempts to reveal the nature of processes involved via computational modeling. Investigations of infant language-appropriate perception are rarely able to obtain more than a few data points per individual because of experimental constraints involved with testing infants. With such sparse data, computational modeling becomes difficult because nearly any model can be made to fit several data points. The explanatory power of such an effort is quite limited. However, greater experimental control offered by animal models can result in finely-grained sampling of effects of experience on perception. With rich data sets like these, it becomes possible to meaningfully yoke performance of an individual subject to the output of a computational model and explore the explanatory power of several different assumptions.

In modeling, the nature of putative learning mechanisms may become clearer. For now, the present approach is not wed to a particular class of learning. As a first step, it is sufficient to recognize that organisms are exquisitely sensitive to regularities of the environment and attempt to discover how much of the structure of language acquisition can be accounted for by learning these regularities. The types of learning involved in this task remain to be discovered. These sensitivities are not likely linked to Skinnerian or operant learning (e.g., Skinner, 1957), but rather may arise from adaptive self-organization (e.g., Kohonen, 1987) or other "unsupervised" means of pattern classification. Data collected from animal models of infant language acquisition are sure to provide a useful corpus with which to explore the plausibility of these possibilities computationally.

Surprisingly little attention has been devoted to examining role of learning or to developing theoretical accounts of learning mechanisms in language acquisition. The approach advocated here aims to remedy this oversight. Yet, suggesting that it is useful to examine general learning mechanisms before postulating innate predisposition does not imply that no innate biases guide language acquisition. If elements of language acquisition can be accounted for by principles of general learning, then there is no need to postulate complex innate constructs, species-specific mechanisms, or computationally-laden alternatives.

Notes

1. Click consonants are produced by formation of suction in the oral cavity followed by an abrupt release of negative pressure (Doke, 1926).
2. For reasons explained more fully elsewhere (Kluender et al., 1998a), the terms "category" and "categorization" are used hesitantly here. These terms have a theoretically-laden history. The present use does not intend to imply any particular internal cognitive representation.
3. Best (1995) makes a similar argument from the theoretical standpoint of direct realism.

References

Best, C. T. 1995. A Direct Realist View of Cross-Language Speech Perception. In W. Strange (Ed.) Speech Perception and Linguistic Experience. York Press: Baltimore, MD.

_____, McRoberts, G., LaFleur, R., & Silver-Isenstadt, J. 1995. Divergent Developmental Patterns for Infants' Perception of Two Non-Native Consonant Contrasts. Infant Behavior and Development 18. 339-350.

_____, McRoberts, G. W., & Sithole, N. N. 1988. The Phonological Basis of Perceptual Loss for Non-Native Contrasts: Maintenance of Discrimination Among Zulu Clicks by English-Speaking Adults and Infants. Journal of Experimental Psychology: Human Perception and Performance 14. 345-360.

Chomsky, N. 1959. Review of Verbal Behavior, by B. F. Skinner. Language 35. 26-58.

Dent, M. L, Brittan-Powell, E. F., Dooling, R. J., & Pierce, A. 1997. Perception of Synthetic /ba/-/wa/ Speech Continuum by Budgerigars (*Melopsittacus undulatus*). Journal of the Acoustical Society of America 102. 1891-1897.

Diehl, R. L. 1991. The Role of Phonetics Within the Study of Language. Phonetica 48. 120-134.
_____, & Kluender, K. R. 1989. On the Objects of Speech Perception. Ecological Psychology 1. 121-144.
Doke, C. M. 1926. The Phonetics of the Zulu Language. In Bantu Studies: Volume 2. University of Witwatersrand Press: Johannesburg.
Dooling, R. J., Best, C. T., & Brown, S. D. 1995. Discrimination of Synthetic Full-Formant and Sinewave /ra-la/ continua by Budgerigars (*Melopsittacus undulatus*) and Zebra Finches (*Taeniopygia guttata*). Journal of the Acoustical Society of America 97. 1839-1846.
Eimas, P. D. 1987. Speech Perception in Early Infancy. In P. Salapatek & L. Cohen (Eds.) Handbook of Infant Perception. Academic Press: Orlando.
Estes, W. K. 1950. Toward a Statistical Theory of Learning. Psychological Review 57. 94-107.
Friederici, A. D., & Wessels, J. M. I. 1993. Phonotactic Knowledge and Its Use in Infant Speech Perception. Perception and Psychophysics 54. 287-295.
Hebb, D. O. 1949. The Organization of Behavior. Wiley: New York.
Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Wright Cassidy, K., Druss, B., & Kennedy, L. 1987. Clauses are Perceptual Units for Young Infants. Cognition 26. 269-286.
Johnson, E. K., Saffran, J. R., Aslin, R. N., & Newport, E. L. (1998). Statistical Learning of Non-Linguistic Stimuli by 8-Month-Olds. International Conference on Infancy Studies, Atlanta, April, 1998.
Jusczyk, P. W., Cutler, A. & Redanz, N. 1993. Preference for the Predominant Stress Patterns of English Words. Child Development 64. 675-687.
_____, Friederici, A. D., Wessels, J., Svenkerud, V. Y., & Jusczyk, A. M. 1993. Infants' Sensitivity to the Sound Patterns of Native Language Words. Journal of Memory and Language 32. 402-420.
_____, Luce, P. A., & Charles-Luce, J. 1994. Infants' Sensitivity to Phonotactic Patterns in the Native Language. Journal of Memory and Language 33. 630-645.
Kluender, K. R., Diehl, R. D., & Killeen, P. R. 1987. Japanese Quail Can Form Phonetic Categories. Science 237. 1195-1197.
_____, & Lotto, A. J. 1994. Effects of First Formant Onset Frequency on [-Voice] Judgements Result from General Auditory Processes not Specific to Humans. Journal of the Acoustical Society of America 95. 1044-1052.
_____, Lotto, A. J., Holt, L. L., & Bloedel, S. L. Accepted Pending Minor Revisions. Role of Experience in Language-Specific Functional Mappings for Vowel Sounds as Inferred from Human, Nonhuman, and Computational Models. Journal of the Acoustical Society of America.
_____, Lotto, A. J., & Holt, L. L. In Press. Contributions of Nonhuman Animal Models to Understanding Human Speech Perception. In S. Greenberg and W. Ainsworth (Eds.) Listening to Speech: An Auditory Perspective. Oxford University Press: New York.
Kohonen, T. 1987. Self-Organization and Associative Memory. Springer-Verlag: Berlin.
Kuhl, P. K. 1983. Perception of Auditory Equivalence Classes for Speech in Early Infancy. Infant Behavior and Development 6. 263-285.
_____, 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.
_____, & Miller, J. D. 1975. Speech Perception by the Chinchilla: Voiced-Voiceless Distinction in Alveolar Plosive Consonants. Science 90. 69-72.
_____, & Miller, J. D. 1978. Speech Perception by the Chinchilla: Identification Functions for Synthetic VOT Stimuli. Journal of the Acoustical Society of America 63. 905-917.
_____, & Padden, D. M. 1983. Enhanced Discriminability at the Phonetic Boundaries for the Place Feature in Macaques. Journal of the Acoustical Society of America 73. 1003-1010.

_____, Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. 1992. Linguistic Experiences Alter Phonetic Perception in Infants by 6 Months of Age. Science 255. 606-608.
Lisker, L., & Abramson, A. S. 1964. A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. Word 20. 384-422.
Lotto, A. J., Kluender, K. R., & Holt, L. L. 1997. Perceptual Compensation for Coarticulation by Japanese Quail (*Coturnix coturnix japonica*). Journal of the Acoustical Society of America 102. 1134-1140.
MacKain, K. S. 1982. Assessing the Role of Experience on Infants' Speech Discrimination. Journal of Child Language 9. 527-542.
Morse, P. A., & Snowdon, C. T. 1975. An Investigation of Categorical Speech Discrimination by Rhesus Monkeys. Perception and Psychophysics 17. 9-16.
Polka, L., & Werker, J. F. 1994. Developmental Changes in Perception of Nonnative Vowel Contrasts. Journal of Experimental Psychology 20. 421-435.
Saffran, J. R., Aslin, R. N., & Newport, E. L. 1996. Statistical Learning by 8-month-olds. Science 274. 1926-1928.
Skinner, B. F. 1957. Verbal Behavior. Appleton-Century-Crofts: New York.
Spence, K. W. 1936. The Nature of Discrimination Learning in Animals. Psychological Review 43. 427-449.
_____, 1937. The Differential Response in Animals to Stimuli Varying Within a Single Dimension. Psychological Review 44. 430-444.
_____, 1952. The Nature of the Response in Discrimination Learning. Psychological Review 59. 89-93.
_____, 1960. Behavior Theory and Learning. Prentice-Hall: Englewood Cliffs, NJ.
Trehub, S. E. 1976. The Discrimination of Foreign Speech Contrasts by Infants and Adults. Child Development 47. 466-472.
Waters, R. A., & Wilson, W. A., Jr. 1976. Speech Perception by Rhesus Monkeys: The Voicing Distinction in Synthesized Labial and Velar Stop Consonants. Perception and Psychophysics 19. 285-289.
Werker, J. F., Gilbert, J. H. V., Humphrey, K., & Tees, R. C. 1981. Developmental Aspects of Cross-Language Speech Perception. Child Development 52. 349-353.
_____, & Lalonde, C. E. 1988. Cross-Language Speech Perception: Initial Capacities and Developmental Change. Developmental Psychology 24. 672-683.
_____, & Logan, J. S. 1985. Cross-Language Evidence for Three Factors in Speech Perception. Perception and Psychophysics 37. 35-44.
_____, & Tees, R. C. 1984a. Cross-Language Speech Perception: Evidence for Perceptual Reorganization During the First Year of Life. Infant Behavior and Development 7. 49-63.
_____, & Tees, R. D. 1984b. Phonemic and Phonetic Factors in Adult Cross-Language Speech Perception. Journal of the Acoustical Society of America 75. 1865-1878.
Young, J. Z. 1962. Why do we have two brains? In V. B. Mountcastle (Ed.) Inter-hemispheric relations and cerebral dominance. The Johns Hopkins University Press: Baltimore.