Revisiting Roger Brown's "Original Word Game":
An Experimental Approach to the Pseudo-Semantic Basis of
Language-Specific Speech Perception in Late Infancy

Reese M. Heitner
Graduate School and University Center
City University of New York
33 West 42nd Street
New York, NY 10036, USA
rheitner@willkie.com

Abstract
Recent experimental research in developmental phonology confirms that within their first year, infants demonstrate language-specific patterns of speech perception discrimination. Data, however, outpaces theory in this line of research. For while it is clear that this transition is related to the particular linguistic environment of the infant, it remains obscure whether mere acoustic exposure to—as opposed to interpretive experience with—the target language is sufficient to support the details of this development. The logical dynamic language-specific allophonic variation imposes on phonological development is integrated into an account of the transition from language-general to language-specific speech perception observed within the first year of infancy. An experiment, reminiscent of Brown's "Original Word Game", but specifically designed to test the "cognitive" hypothesis that 8- to 10-month old subjects can exploit nascent speech-referent pairs to bootstrap phonological development is proposed.

The child passes through words to the phonemic perception of speech.

N. Kh. Shvachkin

Developmental Speech Perception
Intrigued by the robust pace of language acquisition, linguists have proposed a variety of adaptive mechanisms which support, and explain the experimental details of children's linguistic progress. In particular, speech theorists have documented neonates' precocious language-like speech discrimination (Eimas 1975, Kuhl 1987, Werker 1994), while developmental psycholinguists have detailed the assumptions or biases guiding preschoolers' verbal-conceptual maturation (Gelman and Wellman 1991, Keil 1991, Markman 1989). Though developmental constraints are often portrayed as innate, linguistic accelerants need not be specified in the genome. Rather, they may be part of what Brown (1958) dubbed the "Original Word Game". As current experimental research demonstrates, external information provided in the form of word labels can facilitate children's superordinate and atypical object categorization (Gelman and Coley 1990, Waxman and Markov 1995). By contrast, however, the inverse hypothesis that the external linguistic labeling of "basic-level" objects (Rosch et al. 1976) can developmentally support the language-specific phonological reorganization (Best 1994, Jusczyk 1993, Werker and Tees 1992) of infant speech categories has yet to be theoretically isolated and experimentally explored.

Broadly speaking, contemporary investigations into developmental speech perception represent an attempt to synthesize two strains of empirical research that lead in opposite directions. On the one hand, a variety of now-standard experiments in infant speech perception indicate that neonates are innately prepared to discriminate among most, if not all, phonetic speech contrasts which happen not to be phonemic in their native language (Pisoni et al. 1982, Trehub 1976, Werker 1989). In light of this, researchers turned their attention to documenting when this transition from language-general to language-specific speech perception occurs. They soon discovered, at least to their initial surprise, that within their very first year, infants undergo a process of phonological specialization characterized, at least in part, by a language-specific inability to discriminate among various non-phonemic contrasts (Jusczyk 1994, Werker and Tees 1992). As Eimas (1985) concludes,

it would seem that the infant comes into the world biologically endowed with the ability to distinguish and categorize virtually all of the information that is

From:MAICS-99 Proceedings. Copyright © 1999, AAAI (www.aaai.org). All rights reserved.
relevant to the phonetic categorization of natural languages. Experience with the parental language serves to maintain and perhaps enhance those categorizations for which there is an early correspondence. On the other hand, experience will eliminate a category or alter the boundary location between categories in situations where the correspondence is quite disparate (p. 171).

Data, however, outpaces theory in this line of developmental research. For while it is now obvious that this transition is related to the particular linguistic environment of the infant, it remains obscure whether mere acoustic experience with the target language is sufficient, or perhaps interpretive experience with the language is also necessary, to explain the dynamic of this selective perceptual desensitization.

"Linguistic Experience": Acoustic or Interpretive?

To be sure, one reason why these two linguistic variables (i.e., acoustic experience and interpretive experience) are so easily confounded is the obvious point that 8- to 10-month old infants' acoustic and incipient interpretive experiences are quite naturally entangled. Infants typically do not hear human speech in deliberately contrived interpretive vacuums, and it would seem that no (ethical) experiment which attempted to control for this variable by exposing infants only to phonologically correct—but perpetually meaningless—speech can be conducted. Moreover, if the development of language-specific speech perception is indeed at least partially underwritten by pseudo-semantic interpretive information, this argument incurs the empirical burden that approximately 8- to 10-month old infants are not only capable of organizing their environment into (perceptually, functionally, thematically, etc. based) categories, but are also capable of associating particular speech tokens with these non-linguistic categories. Nevertheless, experimental data marshaled from pre-linguistic infant categorization (Eimas 1994, Mandler 1995, Younger 1993), combined with studies investigating infant "cross-modal" speech-object association (Baldwin and Markman 1989, Mandel et al. 1995, Waxman and Markow 1995)—not to mention early infant comprehension studies (Bates et al. 1994, Benedict 1979, Harris et al. 1995, Huttenlocher 1974)—suggest that such speech-object associations might very well be in place. Be that as it may, the causal relationship between this putative interpretive ability and the development of language-specific speech perception has yet to be theoretically established.

Consider, however, how the pseudo-semantic skills of early speech-object associations could potentially induce suppressed discriminability of certain phonetic contrasts, resulting in language-specific speech perception: Insofar as infants attribute some interpretive value (whatever that might be) to some speech utterances, and at least some of these utterances are subject to sub-phonemic variation, then the fact that two phonetically distinct speech tokens are consistently associated with the same (class of) object(s) might "referentially" motivate the phonologically naive infant to collapse those phonetic differences that, as it were, simply do not make a difference and begin to perceive them in accordance with the target phonology. No doubt this is why Walley (1993) observes that the language-specific changes in infants' speech perception documented by Werker and colleagues occur "perilously close to the nine-month point often cited as the age at which infants first begin to comprehend words" (p. 173).

"Free" Allophonic Variation

In addition, then, to capitalizing on the suggestive but largely circumstantial evidence provided by empirical studies of early infant speech-object association, the logical dynamic sub-phonemic variation imposes on phonological development must also be considered when evaluating the proposal that language-specific speech perception might be at least partially driven by semantic interpretive cues. In particular, accounts of phonological development which implicitly assume that language-specific speech perception can emerge in an interpretive vacuum are arguably incomplete on purely methodological grounds. To see why, consider the fact two languages may draw upon the same phonetic resources but nevertheless instantiate different phonological systems whereby the same pair of distinct phones is phonemically contrastive in one language but phonemically irrelevant in the other (Sommerstein 1977). This phenomenon—known as language-specific allophonic variation—can result in suppressed discriminability of phonetically distinct but allophonically related contrasts (Pegg and Werker 1997).

The issue, then, reduces to whether, say an English-learning infant could, at least in principle, eventually learn to correctly categorize English speech tokens merely by listening to extended bouts of English heard exclusively over the telephone or radio. Yet because this auditory stimulus may very well include speech tokens which are phonetically distinct, but (at least in English) phonemically equivalent (e.g., in utterance-final positions /p/ may be aspirated or unaspirated and /d/ may be released or unreleased), there is no (acoustic) reason for the infant to treat them as allophonic variants, nonetheless desensitize to the contrast. After all, these very same allophonic contrasts could very well be full-blown phonemic contrasts just as much as aspirated and unaspirated /p/--notationally [p'] and [p]—actually are phonemically contrastive in Hindi. This is why [pal] (translated as "taking care of") and [pal] ("edge of knife") are recognized as two different words in Hindi (MacKain 1985,
MacKain and Stern 1985), while in English they are one and the same word.

Of course, the point of this example is that both the Hindi-learning infant and the English-learning infant are exposed to [pʰ] and [p]. And the question is how is the infant supposed to figure out whether they are distinctive phonemes, as in Hindi, or allophonic variants, as in English? Clearly, by appreciating the semantic difference between "taking care of" and "edge of knife" and the semantic equivalence between, say [stop] and [stopʰ], the infant is positioned to induce the communicative, and hence, language-specific phonological (ir)relevance of this phonetic contrast. Insofar as these two phones are--ignoring phonotactic regularities for the moment--distributed randomly (if not also equally) in the speech stream of both English and Hindi, no amount of mere auditory exposure is sufficient to induce the phonological reorganization of the phonetic contrast simply because both contrasts appear in both languages. The only feature which may distinguish such contrasts is that they are semantically pregnant in one language but meaningless in another.

While some researchers have skeptically speculated about the role interpretive cues might play in inspiring language-specific speech perception in infancy (Jusczyk 1985, Kuhl et al. 1992, Macnamara 1982, Vihman 1996, Werker and Tees 1984), only MacKain (MacKain 1982, MacKain and Stern 1985) has been particularly forceful in highlighting the role semantic-like information, in contrast to mere acoustic exposure, could play in phonological development:

"Experience listening to the speech of a particular linguistic community will not induce discrimination of a language-specific contrast, shift a boundary to correspond to that of the language user, or attenuate discrimination of a distinction that may be discriminable by virtue of its psychoacoustic characteristic until infants are aware that certain sounds contrast to convey differences in meaning (MacKain and Stern 1985, p. 31).

To be sure, MacKain's case is somewhat overstated. To the extent that some phonetic contrasts will not appear in the infant's target language simply because every phone does not appear in every language, the infant may develop language-specific patterns of speech perception on the basis of what amounts to the phonological equivalent of negative evidence. Moreover, insofar as allophonic variation is predictable on the basis of language-specific patterns governing the phonetic co-articulation of the same phoneme in different contexts (Jusczyk 1993, Jusczyk and Aslin 1995), then the phonology of the language can, at least in part, be inferred by more or less passively tabulating these statistical phonotactic regularities. As such, the answer to MacKain's (1988) question: "Can exposure to the speech of their native language influence infant's perception of phonetic contrasts prior to the onset of symbolic functioning?" (p. 55) is clearly 'yes'. On the other hand, as MacKain (1982) emphasizes and Jusczyk (1993) seems to implicitly recognize, the presence of "free" allophonic variation presents a formidable challenge to the establishment of language-specific speech categories simply because it undermines the "well-behaved" phonotactic regularities of co-articulation distributionally available to the infant in the speech stream.

It is for these reasons, then, that it would seem that the full phonological structure of language cannot be induced merely through asemantic acoustic exposure. But nor can we, for obvious reasons, attempt to empirically confirm this reasoning simply by exposing infants to exclusively meaningless speech. Instead, we must concentrate on whether, at least under favorable conditions, interpretive-based information can be causally implicated in the acceleration of 8- to 10-month old infants' perceptual desensitivities to phonetically distinct speech tokens. So rather than underrate the role (acoustically accessible) phonotactically regularities play in supporting phonological development, we must look for evidence that, in addition to relying on information intrinsic to the speech pattern of their target language, infants may also capitalize on pseudo-semantic information extrinsic to what the speech patterns are about.

Revisiting the "Original Word Game"

With this in mind, it should now be clearer how to experimentally substantiate the role nascent referential categories might play in underwriting the development of language-specific (allophonic) speech perception. In fact, it is merely a version of Brown's "Original Word Game". Using the "Shvachkin-Garnica" word-object training procedure originally devised by Shvachkin (1973) and Garnica (1973) for testing phonemic speech discrimination in preschoolers--recently adapted by Werker and Pegg (1992) for infants--but inverting it to test for phonemic speech equivalencies, the hypothesis that infants use non-linguistic categories to help convert universal speech perception into language-specific speech perception can be empirically tested. Indeed, simply by systematically associating both tokens of a phonetically contrastive minimal "word" pair with either the same referential category to encourage suppressed discrimination, or uniquely associating each phonetically contrastive speech token with dissimilar referential categories to discourage suppressed discrimination, infant phonological development should be susceptible to experimental manipulation. (Baseline object similarity/dissimilarity is pre-determined by using standard categorization tasks, e.g., habituation-dishabituation, sequential touching, suitable for 8- to 10-month old infants (Younger 1993), while a mere 15 minutes (5 minute intervals over three days) of laboratory exposure has been shown to be
effective in influencing infant responses to various phonetic stimuli (Kuhl and Meltzoff 1997).

In other words, once infant speech perception is recognized to consist in the development of language-specific desensitivities, the basic experimental task designed to explore this phenomena must be revised to reflect this change in theoretical orientation. Rather than testing infants for language-specific phonetic discrimination, infants must be tested for phonetic attenuation in order to demonstrate the role interpretive information might play in this process. This is why Werker’s (Werker and Pegg 1992, Stager and Werker 1997) hesitancy to grant 10- to 12-month infants a language-specific phonological system because they do not interpretively respond to words they do understand differently from phonemically contrastive nonsense words they do not understand (e.g., ‘dog’ and ‘bog’) is noteworthy but misplaced. Pace Werker, it is not the case that “the strongest evidence for the functional (phonological) use of perceptual categories would be provided if it could be shown that infants differentiate real words on the basis of minimal pair contrasts” (Werker and Pegg 1992, p. 300 italics added), simply because, as reviewed above, pre-linguistic infants are essentially born with the ability to discriminate subtle minimal pair (phonetic) contrasts. Instead, the relevant “litmus test” for the early “functional” appearance of language-specific phonological processing is that infants be able to apply their evolving perceptual sensitivities to the task of NOT distinguishing between speech tokens on the basis of minimal pair allophonic contrasts. What is diagnostic of a semantically-informed language-specific phonological system in late infancy, then, is not the interpretive ability to discriminate between, say ‘dog’ and ‘bog’, but the interpretive ability not to discriminate between ‘dog’ as it might be phonetically realized say, with and without prevoicing.

This paradigm then, rather than indirectly correlating a constellation of cognitive skills (e.g., object categorization, mean ends, arbitrary association) with language-specific speech perception as per Werker (Lalonde and Werker 1995), attempts to directly examine the influence of linguistically-labeled categories on the development of speech perception. As Walley (1993) reminds us,

In light of the rather slim three-month difference between current estimates for the onset of language-specific speech perception and for word comprehension, it would seem essential in future research to assess directly those core cognitive/semantic abilities that infants are claimed to lack, in addition to demonstrating that they possess a certain perceptual ability at a particular age. That is, it will be important to directly link or unlink, as the case may be, performance across different linguistic subdomains within subjects. This approach would yield the strongest and most convincing data for evaluating different theoretical accounts of phonetic development (p. 174)

In our experiment, the two minimal “word” pair phonetic contrast training conditions are designed to be identical but for the manipulation of the similarity of the referential categories. In the first condition, the same category is used so as to suppress discriminability of the phonetically contrastive “word” label(s), while in the second condition, distinct referential categories are used so as to preserve the phonetic discriminability of the two phonetically contrastive "word" labels. For this reason, if, after training, statistically significant differences in discrimination are observed when the subjects are tested on novel minimal word pairs utilizing the very same phonetic contrast, then it would seem the variable of non-linguistic object categorization can influence speech perception—just as Brown presciently suggested some thirty years ago.

After all, recall that the "original word game" experiments Brown (1958) reported were designed to show the effects of speech perception on object categorization, as well as the effects of object categorization on speech perception. With respect to the former dynamic, remember that the native Navaho speakers "heard" a total of four words whereas the English speakers "heard" only two words, and this difference in language-specific speech perception influenced the object categorization behavior of the (adult) subjects. While successfully demonstrating, as he put it, "speech categories operating as a guide to referent categories", Brown succeeding, albeit crudely by today’s standards, in implicitly demonstrating the more general effect of linguistic labeling on object categorization that has since become a major focal point of current developmental psycholinguistic research (Davidson and Gelman 1990, Waxman and Markow 1995).

But by also running the experiment essentially in reverse, Brown discovered that non-linguistic (perceptually-based) object categories could serve as a guide to the categorization of the experimenter’s referential speech. By deliberately biasing the initial perceptual categorization of the objects available to subjects, Brown found English subjects could be prompted to parse the speech tokens according to non-native contrasts. As Brown (1958) concluded, "This result demonstrates a facet of the Word Game that we have not yet discussed. It is evidently possible for nonlinguistic reality to serve as a guide to the categorization of speech. The isomorphic relationship can be useful in either direction. An inescapable visual difference leads us to look for a speech difference" (p. 216). Of course, the counterpart of this last claim—namely that "inescapable similarities" among objects may lead us, and in particular 8- to 10-month old infants, to "look for" speech equivalencies among phonetically distinct phones—is precisely the thesis explored in this paper. And while collapsing phonetic contrasts on the basis of interpretive identity is obviously fallible—infants would be, as Macnamara (1982) points out,
"phonological anarchists" if they attempted to phonemically equate 'small' and 'tittle' merely because they mean very nearly the same thing—this source information could serve as a valuable developmental mechanism supporting the appearance of language-specific "perceptual accents" in late infancy and deserves experimental scrutiny.

References


