The Acquisition of Language Specific Phonetic Categories in Infancy

Janet F. Werker
Department of Psychology, University of British Columbia
E-mail: jwerker@psych.ubc.ca

ABSTRACT

Maye, Werker, and Gerken [1] have recently shown that distributional learning may allow tuning of phonetic categories during the first year of life. This paper explores the power and potential limits of such a learning mechanism. Burns, Werker & McVie [2] showed the pattern of age-related change in consonant perception to be different for bilingual (French/English) than monolingual infants, with bilingual infants ultimately retaining sensitivity to the voicing distinctions from both of their native languages. The potential role of distributional learning in accounting for this difference will be discussed. But can distributional learning account for phonological and lexical acquisition? The final portion of this paper reviews recent research addressing this question.

1. INTRODUCTION

By their first birthday, infants are exquisitely sensitive to just those sound differences that are important in the language they are learning. In contrast to the rather broad-based phonetic discrimination shown by infants in the first half year of life, by 10-12 months of age infants show selective discrimination of only those minimally different consonants that are used contrastively in their native language. We first showed this by testing English and Hindi adults, as well as English- and Hindi-learning infants on their ability to discriminate syllables differing in the Hindi retroflex vs. dental “d” sounds, the Hindi voiceless aspirated vs. breathy voiced (two differences that are contrastive in Hindi) syllables and the Interior Salish glottalized velar vs. uvular /k’i/-/q’i/ syllables. Although English-learning infants of 6-8 months as well as L1 (Hindi or Nthlakampx) adults and older infants could discriminate these non-English distinctions, English adults and English-learning infants of 10-12 months failed [3].

More recently, Burns, McVie, and I [2] have documented the pattern of change in perception of the voicing distinction in infants growing up in a French, English, or bilingual French-English environment. Infants were tested on their ability to discriminate the French [b] vs. [p] and the English [p] vs. [p^h] distinctions. We found that by the end of the first year of life, English and French monolingual infants discriminate their native voicing boundary better than the non-native boundary, whereas bilingual learning infants of 10-12 months of age seem to show evidence of discriminating neither. By 14-17 months, however, the bilingual infants follow two patterns: approximately half of them behave as either monolingual English or monolingual French infants, and the other half show sensitivity to both the French and the English boundaries. Similar findings have been reported with Spanish-Catalan learning infants’ perception of Catalan vowels. In this case, the Spanish-Catalan bilingual infants of 4-months show sensitivity to both the Spanish and Catalan vowels, but go through a period at around 8-months during which they fail to discriminate Catalan vowels, only to have evidence of discrimination reappear toward the end of the first year of life [4].

In the years following the initial work, our research focused on identifying the underlying mechanism that leads to changes in cross language speech perception. An early hypothesis was that knowledge of words might be the mechanism that allows infants to figure out what information to attend to and what to ignore. The reasoning was that in learning how to map sound on to meaning, a child would need to figure out which differences in pronunciation are criterial, and which are just accidental.

2. TESTING WORD LEARNING

To address the question of whether it is the establishment of phonemic categories that leads to the change in speech perception in infancy, we developed a procedure for testing associative word learning in infancy [5]. In this procedure,

 ISBN 1-876346-48-5

infants are familiarized to two word-object pairings, and then tested on their ability to detect a switch in the pairing. For example, they are familiarized to Word A with moving Object a, and Word B with moving Object b, and then tested on a Same Trial (Aa or Bb) and a Switch Trial (Ab or Ba). By 14 months - but not at 8, 10, or 12 months - infants can learn the association between two words and two moving objects when the words used are phonetically dissimilar (e.g., 'lif' and 'neem').

In a follow-up series of studies, Stager and Werker [6] found that although infants of 14months can learn to map phonetically dissimilar words on to two different objects, they fail to map phonetically similar words such as "bih" and "dih" on to two different objects, even though they can discriminate this b/d phonetic difference in a simple discrimination task. Pater, Stager, and Werker [7] replicated this finding with well-formed words ("bin" vs. "din") and a place ("bin" vs. "pin") as well as a place-plus-voicing ("pin" vs. "din") distinction. It is not until they are 17months of age, or have a particularly large vocabulary at 14 months that infants can learn new words that differ in only a single phonetic feature [8]. Of interest, infants of 14months can use the same bilabial vs. alveolar distinction when performing an identical Switch task with already known words such as "ball" and "doll" [9, 10]. The fact that working knowledge of this phonological opposition in a well-known pair was not sufficient to allow use of this same information in learning new words was the final piece of data refuting our original hypothesis.

3. STATISTICAL LEARNING

Having established that word learning is likely the beneficiary of rather than the engine driving the restructuring of phonetic categories, we went back to the drawing board. In searching for a mechanism that might lead to the restructuring of phonetic categories prior to the acquisition of a lexicon, we turned to the broader literature showing that in addition to becoming sensitive to the phonetic categories of the native language during the first year of life, infants also become sensitive to many other properties of the native language such as phonotactics (the rules for which sequences of sounds can occur together). In learning about these properties, infants seem to use statistical information in the input [11], with the hypothesis being that they calculate statistics about the frequency of letter-to-letter transitions. In an elegant series of studies, Saffran and colleagues [12] showed that infants are exquisitely sensitive to statistical regularities in the input and are able to use statistical information (in her case, transitional probabilities) to segment words from an ongoing stream of speech.

We asked whether statistical regularities in the input might similarly guide the restructuring of phonetic categories. Certainly this hypothesis is a central tenet of Kuhl's perceptual magnet theory [13], and is entirely consistent with Best's Perceptual Assimilation Model [14] wherein the most difficult non-native contrasts are those in which the two non-native phones are thought to be assimilated to an intermediate category in the native language. It is also consistent with recent work by Anderson and Morgan [15] showing that the influence of language-specific experience is evident at an earlier age for non-native phones for which there are highly frequent intermediate categories in the native language. Yet, until our recent work, no one had empirically examined whether or not differential exposure, in a laboratory setting, can change phonetic perception.

In Maye, Werker, and Gerken [1] we manipulated distributional information in the input and measured its impact on phonetic categories. We familiarized infants 6-8 months of age to 8 stimuli ranging from voiced unaspirated to voiceless unaspirated stop. Probably because there was no aspiration, all stimuli were identified by English adults as exemplars of "d". In earlier work, Pegg and Werker [16] had shown that infants of 6-8 months of age can discriminate voiced from voiceless unaspirated stops, whereas English-learning infants of 10-12 months of age have more difficulty. In the Maye, et. al. work we asked whether differential exposure would change this pattern in the younger infants.

The infants of 6-8 months of age were divided into two groups. All infants were presented with at least 4 exemplars of each stimulus along the continuum. The bimodal input group of infants heard more instances (16) of exemplars 2 and 7, corresponding to a language in which both a voiced and a voiceless unaspirated category might exist in the input.

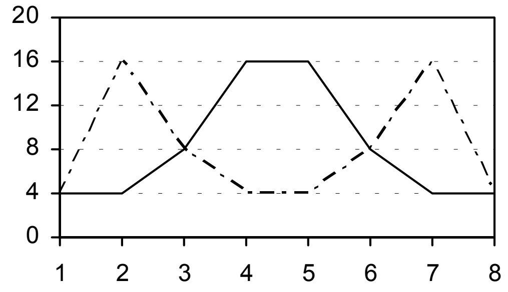

Figure 1: Number of instances presented during the familiarization phase (ordinate) for each stimulus along the continuum (abscissa).

The monomodal group of infants heard 16 instances of exemplars 4 and 5, corresponding to a language like English where the variability in the input in initial position words would likely vary around a central set of values. Critically, all infants heard the same number of exposure to stimuli 1 and 8. Following only 2.4 minutes of familiarization we tested infants on their ability to discriminate the endpoints of the continuum, using the logic that those infants who had received bimodal input may have used that distributional information to divide the continuum into two categories whereas those who had received only monomodal input should treat the continuum all as one category. Our predictions were upheld. Infants in the bimodal, but not the monomodal group were able to discriminate Stimuli 1 vs. 8 following familiarization. This verifies that statistical learning can indeed be used to set up native language perceptual categories in the infancy period.

A distributional learning account provides a potential explanation for the pattern of results seen with bilingual learning infants as well. Specifically, it would predict that performance on phonetic distinction would be based on the relative distribution of those distinctions in the input the child is receiving. A child in a balanced bilingual environment in which both distinctions (e.g. the French b -p and the English p- p^h) are used equally often in both languages, would ultimately retain both distinctions, but might need sufficient exposure to both to show this sensitivity. However, if one language were more common than the other in the input, or if one contrast were more common in one language than in the other, there may be a period in development in which sensitivity to the contrast in only one of the two native languages might be apparent.

4. REMAINING QUESTIONS

Two important questions are still unanswered. 1) Why is it that infants seem not to use their native language phonetic categories in the earliest stages of word learning even though the phonetic categories are well-established? 2) What allows them to be able to use their language-specific phonetic categories a few months later to direct word learning?

In attempting to answer the first of these questions, we suggest that the task of simply linking words with objects in an arbitrary relation is so computationally intensive for novice word learners that the attentional resources are not available to utilize the fine phonetic detail in new words [6, 8,17]. Although this is still a hypothesis, it is supported by recent research. Swingley and Aslin [18] have shown there to be evidence in on-line word recognition tasks that the distinctive information in minimally different words is picked up at some point in the processing stream. We have shown that infants of 14 months who are more accomplished word learners as indicated by larger vocabularies, are able to learn phonetically similar words [8], and that infants of 14months can distinguish phonetically similar words that they already know well when tested in the identical Switch task [9]. For similar results with slightly older infants, see Bailey and Plunkett [19]. All of these results are consistent with the hypothesis that it is a shortage of processing resources that interferes with learning phonetically similar words at 14 months.

An extension of the processing limitation hypothesis suggests that once infants become more accomplished word learners, they have the attentional resources available to utilize the phonetic detail in new words in a word-object associative learning task [6, 16]. It may be, however, that more than just processing capacity contributes to this success. We [6] and others [20, 19] have suggested that the achievement of a criterial vocabulary size may also be an essential ingredient to rapidly learning the full phonetic specification of new words. Others have suggested that it is neighborhood density [21]. Recently, Suzanne Curtin and I have developed a model that pulls these three hypotheses together [22]. In the PAUSE: mir model we suggest that

ISBN 1-876346-48-5 © 2003 UAB

prior to a full referential insight, word learning is very much an associative process whereby word forms need to be linked with objects in what is, for the infant, a computationally intensive task. Given the processing load, mistakes might be made in the details picked up in either the object or the word [6]. Initially these word-object pairings are stored as individual exemplars, with little generalizability (explaining why infants in Fennell & Werker, [9] can distinguish a bilabial from an alveolar in a well-known word but not use the same detail in a new word-object pairing). These exemplars do, however, cluster in multidimensional representational space with other word-object pairings that are similar along various dimensions (see Pater, Stager, & Werker, [7] for an optimality theory analysis). With the achievement of a sufficient number of pairings in which the words cluster according to language-specific phonetic properties, higher order regularities emerge. We suggest that as these higher order regularities are solidified, they come to function as "phonemes". Once in place, these "phonemes" can begin to direct information pick-up during word learning, and ease the attentional load of detecting and using the criterial detail when learning a new word-object link.

5. CONCLUSION

In summary, a series of studies are reviewed exploring the link between speech perception and word learning. These studies demonstrate that during the first year of life, infants' phonetic categories change to reflect the properties of the ambient language. Statistical learning, via sensitivity to distributional regularities in the input, could be the mechanism that allows the prelinguistic infant to establish language-specific phonetic categories. Initially, these categories are not fully used in directing information pick-up in word learning. However, as the infant establishes a sufficiently large vocabulary, higher order regularities emerge among the many words with similar phonetic features in the infant's repertoire. This coalescence of regularities then functions as a phoneme, helping to direct attention to just that acoustic/phonetic variation in new words that is essential for fully integrating the words into a lexical-phonological system.

REFERENCES

[1] J. Maye, J. Werker, & L. Gerken, "Infant sensitivity to distributional information can affect phonetic discrimination", *Cognition, 82*(3), B101-B111, 2002.

[2] T. Burns, J. Werker, & K. McVie, "Development of phonetic categories in infants raised in bilingual and monolingual environments", in *Proceedings of the 27th annual Boston University Conference on Language Development*. Cascadila Press, in press, 2003.

[3] L. Bosch and N. Sebastian-Galles, "Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life", *Language and Speech*, in press.

[4] J. Werker and R. Tees, "Cross-language speech perception: Evidence for perceptual reorganization during the first year of life", *Infant Behavior and Development, 7*(1), 49-63, 1984.

[5] J. Werker, L. Cohen, V. Lloyd, M. Casasola, & C. Stager, "Acquisition of word-object associations by 14-month-old infants", *Developmental Psychology, 34*(6), 1289-1309, 1998.

[6] C. Stager and J. Werker, "Infants listen for more phonetic detail in speech perception than in word-learning tasks", *Nature, 388*(6640), 381-382, 1997.

[7] J. Pater, C. Stager, & J. Werker, "The lexical acquisition of phonological contrasts", *Language*, revision submitted.

[8] J. Werker, C. Fennell, K. Corcoran, & C. Stager, "Infants' ability to learn phonetically similar words: Effects of age and vocabulary size", *Infancy. 3*(1), 1-30, 2002.

[9] C. Fennell and J. Werker, "Early word learners ability to access phonetic detail in well-known words", *Language and Speech*, 2003, in press.

[10] D. Swingley, "Phonetic detail in the developing lexicon", *Language and Speech*, 2003, in press.

[11] P. Jusczyk and P. Luce, “Infants' sensitivity to phonotactic patterns in the native language”, *Journal of Memory and Language.* 33, 630-645, 1994.

[12] J. Saffran, R. Aslin, & E. Newport, “Statistical learning by 8-month-old infants”, *Science.* 274(5294), 1926-1928, 1996.

[13] P. Kuhl, “Innate predispositions and the effects of experience in speech perception: The native language magnet theory”, in *Developmental neurocognition: Speech and face processing in the first year of life*, B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage & J. Morton, Eds, pp. 259-274. Kluwer Academic Publishers, 1993.

[14] C. Best, “Learning to Perceive the Sound Pattern of English”, *Haskins Laboratories Status Report on Speech Research,* 114, 31-79, 1993.

[15] J. Anderson and J. Morgan, “A statistical basis for speech sound discrimination”, *Language and Speech*, in press.

[16] J. Pegg and J. Werker, “Adult and infant perception of two English phones”, *Journal of the Acoustical Society of America, 102*(6), 3742-3753, 1997.

[17] J. Werker and C. Fennell, “From listening to sounds to listening to words: Early steps in word learning”, in *Weaving a lexicon*, G. Hall & S. Waxman, Eds, MIT Press, in press, to be published January 2004.

[18] D. Swingley and R. Aslin, “Lexical neighbourhoods and the word-form representations of 14-month-olds”, *Psychological Science*, 13, 480-484, 2002.

[19] T. Bailey and K. Plunkett, “Phonological specificity in early words”, *Cognitive Development, 17*(2), 1267-1284, 2002.

[20] M. Beckman and J. Edwards, "The ontogeny of phonological categories and the primacy of lexical learning in linguistic development”, *Child Development, 71*(1): 240-249, 2000.

[21] G. Hollich, P. Jusczyk, & P. Luce, “Lexical neighborhood effects in 17-month-old word learning”, in *Proceedings of the 26th Annual Boston University Conference on Language Development*. Boston, MA: Cascadilla Press, 2002.

[22] J. Werker and S. Curtin, “The PAUSE: mir model of infant speech processing,” in *The Handbook of Speech Perception,* D.P. Pisoni and R. Remez, Eds, MIT Press, to appear.

 ISBN 1-876346-48-5 © 2003 UAB