The Level of Detail in Infants' Word Learning

Rochelle S. Newman

University of Maryland, College Park

ABSTRACT—*This article summarizes recent findings on infant word learning and recognition. Infants initially store very detailed representations of words, including details that are not truly necessary for word recognition. As they are exposed to more varied productions of words, they develop more sophisticated knowledge about which details are important, and streamline their representations, allowing them to better recognize words across different contexts, speakers, and environments.*

KEYWORDS—*word learning; infant; word recognition; talker variability; generalization*

Learning and recognizing the sound patterns of new words is a critical part of learning language, and some of the most exciting recent work on this topic has examined how much detail infants store about words they are learning. Even a simple spoken phrase such as "See the cat!" contains a wealth of information, including information about the phonetic and lexical makeup of the utterance, the identity of the talker, and the talker's tone of voice and emotional state.

There are two main schools of thought on how much of that information is stored by infants in their representations. Infants enter the world unaware of what information is important in their particular language. To avoid missing potentially critical information, infants might attempt to store all information, including not only precise details of a word's phonetic structure but also details of the talker's particular pronunciation and tone of voice, and then learn to sort out the important information at a later time. However, compared to adults, infants have limited memory and processing skills, making detailed storage potentially difficult. Some scholars argue that this means infants begin with a very broad approach, devoting what resources they have to gaining the "gist" of the utterance, which they could later refine with further linguistic experience.

Address correspondence to Rochelle Newman, Department of Hearing and Speech Sciences, 0100 Lefrak Hall, University of Maryland, College Park, MD 20742; e-mail: rnewman@hesp.umd.edu.

Current research is leading to the consensus that infants err on the side of storing too much information rather than storing too little, at least from an adult native speaker's point of view. This overspecificity results in infants' initial failure to generalize across different exemplars of the same word (i.e., productions spoken by different talkers). As infants gain experience with their language, they learn what information is critical to word meanings (the sound of the word itself) and what is not (who is speaking); this allows them to generalize across different word exemplars and to recognize words on the basis of partial or noisy productions. Evidence supporting this argument comes from work on both infants' learning of new word forms and their subsequent recognition of words to which they've already been exposed.

LEVEL OF DETAIL IN INFANT WORD LEARNING AND RECOGNITION

To learn a new word, an infant must store enough information to distinguish that word from other words. Because infants' vocabularies are relatively small, they might not need to store as much detail as would an adult to keep words distinct from one another. Although infants could succeed with only vague representations of new words, it appears that they actually maintain a great deal of acoustic detail in their early representations. After being familiarized with the word *cup*, infants listen longer to stories about cups than they do to stories about dogs, demonstrating that they have stored the word *cup* in memory. They will not listen longer to stories about "cuts" or "tups," despite the similarity between these words and the word with which they were just familiarized (cf. Jusczyk & Aslin, 1995). This implies that infants do not simply learn vague word patterns.

Infants' attention to detail is not limited to the specifics necessary for appropriate recognition, and in fact it can impede their ability to recognize the same word in different situations. After being familiarized with *cup* by a female speaker, 7.5-month-old infants do not listen longer to that word when it is spoken by a male (Houston & Jusczyk, 2000). Likewise, if the word was initially spoken in a happy voice, infants will not listen longer to

the word when it is spoken in a neutral voice (Singh, Morgan, & White, 2004). The fact that infants appear very hesitant to generalize across different exemplars suggests that infants' word learning might actually capture too much detail, being tied to specifics of the signal that are actually irrelevant. At 10.5 months, infants are better able to generalize across both gender and affect, suggesting that this overspecificity is a greater issue for younger infants.

Yet even older infants remain tied to details of the signal that are not strictly necessary. Hollich (2006) familiarized 23-month-old infants with two new words, each spoken by a different talker. He subsequently taught them meanings for those two words. When the teaching phase occurred in the same voice as had been heard originally, infants successfully learned the words. When training occurred in a different voice than the original familiarization, however, infants failed to learn the words, demonstrating an overdependence on contextual information that is unnecessary (even detrimental) for the task at hand.

The same pattern is visible when infants process words they already know. In studies exploring word recognition, children might see two television screens showing different objects, such as a cat and a dog (cf. Swingley & Aslin, 2002). The auditory stimulus consists of either correct pronunciation of the objects' names or mispronunciations of them. If children's representations were vague, they should not notice mispronunciations, and should look at the appropriate object for both correctly produced and misproduced sequences. But in fact, 14-month-old children do not identify words as well when they are mispronounced.

Although infants notice mismatching phonetic information, they do not actually need access to a complete word form in order to recognize it, at least among a limited number of alternatives. By 18 months of age, infants presented with only a word's beginning (such as *bay-* from *baby*) will still look at the appropriate object (Fernald, Swingley, & Pinto, 2001). Similarly, results from my laboratory (Newman, 2006) showed that 2-year-olds looked longer at the appropriate object when part of a word was either missing or masked by noise. Yet, although toddlers may realize that the part-word *dino-* is more likely to be *dinosaur* than *lobster*, they show much greater uncertainty in their choice than do older children, even though they only have two objects to choose from (Newman, 2006). Even for well-known words, children appear to need more of the word's information to be intact than do adult listeners—another sign of overreliance on acoustic information in the input.

Thus, regardless of whether the task involves learning word forms alone, mapping forms onto meaning, or recognizing already-known words, children's representations contain a great deal of detail. In some cases, their representations include too much detail, resulting in, from an adult's point of view, too narrow a representation (although such information could be useful for other types of learning).

These findings do not mean that infants are unsophisticated at speech processing. Quite the contrary. Although 7.5-month-old infants appear to have difficulty ignoring irrelevant information about a talker in the word-familiarization task described earlier, infants of the same age do not have difficulty separating out irrelevant background noise. When familiarized with novel words in the presence of noise, infants nonetheless recognize those words later when presented without such noise (Newman & Jusczyk, 1996). Thus, infants seem to distinguish between acoustic information that is contained within the word itself and acoustic information that occurs at the same time but originates from another source. This implies that they have solved the "figure–ground" problem for the auditory domain. It also implies that infants are not simply storing wholesale the entire acoustic signal that hits their eardrums (that is, they are not behaving like a tape recorder); their maintenance of information about talker identity and tone of voice is not because they failed to process the signal (or an indication that they simply require more matching features than adults to recognize a word). Rather, infants recognize that some forms of acoustic information (such as background noise) are not important, even though other forms (e.g., talker identity) might be important and thus worth storing.

THE MEMORY PROBLEM

Infants have far more limited memory and processing skills than do adults, and this can prevent their learning acoustic details when faced with more challenging learning situations (Stager & Werker, 1997). Studies demonstrating this have used a paradigm known as the switch task: Infants are trained to link two new words with two new objects and are tested on their ability to recognize when these mappings have been switched. Infants appear surprised (they look longer) when the familiar words are mapped onto the wrong objects if the two words are very dissimilar (such as *lif* and *neem*); when the words are more similar (*bih* and *dih*), 14-month-old infants do not appear to recognize that a switch has occurred. This observation suggests that they failed to notice the difference between the word they initially learned and the new, similar alternative.

The reason for this early failure appears to be the difficulty of the task itself, relative to the infants' cognitive abilities. Unlike the word-learning studies described earlier, the switch task requires not only that infants learn a new word form but also that they simultaneously learn a new concept and the mapping between concept and form. When they know the concepts in advance (Fennell & Werker, 2004), easing cognitive demands, infants succeed at storing detailed information. Likewise, 17-month-old infants recognize a switch in the same procedure at which 14-month-old infants fail. Although children do store detailed representations of words, doing so requires extensive cognitive resources, which are unavailable when the child is faced with competing demands (such as simultaneously learning the words' meanings).

AN EXPLANATION

The benefits of storing such acoustic detail might outweigh the disadvantages. When infants enter the world, they cannot know which sources of information will be most important in their language. Although some sources of information (such as background noise) are unlikely ever to be relevant, other sources of information might be useful in some languages. For example, although tone of voice (as indicated by voice pitch) is not a necessary part of word representations in English, this information is a critical component of lexical representations in languages such as Mandarin. Indeed, the majority of the world's languages are tonal languages (Yip, 2002), in which the same phonetic sequence can have different meanings depending on voice pitch. Rather than risk missing critical information, infants store all information that might conceivably be important, despite the demands this places on their cognitive resources.

This focus on the details in the signal transcends language domains. Children have been argued to be similarly context dependent in the domain of verb learning, failing to extend words to new situations and agents (cf. Naigles & Hoff, 2006). Hollich, Golinkoff, and Hirsh-Pasek (2007) suggest that children may initially associate a novel label not to an object but to the entire scene, including background details. In the domain of grammar acquisition, linguists have long posited the existence of the subset principle (Berwick, 1985), in which learners map input data onto the most conservative set of rules conforming to the data (but see Braine & Brooks, 1995). The rationale for these narrow mappings is that further examples will allow a learner to recover from too specific a grammar or too narrow a lexical interpretation, whereas only explicit correction can allow recovery once an overly general grammar or lexical interpretation has been posited. (Once a grammar is too broad, no examples in the ambient language will contradict it.)

Thus, the notion that children tend not to generalize but instead tend to remain committed to details of the input seems to apply to more than just early word-form learning. It may be a more general processing approach, allowing the infant to postpone making decisions as to which information is critical until further input is available.

SOLVING THE GENERALIZATION PROBLEM

To fully comprehend language, infants must learn to ignore perceptible but irrelevant information such as tone of voice and to recognize words spoken by a variety of talkers. Variability in the input helps infants recognize which acoustic properties are important and which can be ignored. When an infant is familiar with a word spoken only by a particular talker, or in a particular tone of voice, the word's representation is tied to that talker/tone of voice. However, if the infant hears the same word spoken by multiple talkers, in multiple tones of voice, the child learns that these other factors are irrelevant, and the representation becomes less tied to those details. When exposed to repeated instances of *bug* by a single talker, 2-month-old infants will later react to changes either in the syllable (*dug*) or the talker. When familiarized with instances by six different talkers, infants later react to changes in the syllable but not to new talkers (Jusczyk, Pisoni, & Mullennix, 1992). Similarly, familiarization with words spoken with different emotional tones (Singh, Bortfeld, & Morgan, 2002) results in greater generalization, and sentence variability allows infants to capture higher-order grammatical relationships (Gómez, 2002). Across a range of language domains, when exposure is more varied, infants focus less attention on the specific details of the input and instead begin to abstract across exemplars, focusing on areas of commonality.

FUTURE DIRECTIONS

As children gain more linguistic experience, they begin to learn which information is most critical for distinguishing words (cf. Werker & Curtin, 2005). Our next task is to understand how such developmental change comes about. It seems plausible that variability is important for infants to learn the appropriate level of representational detail, and that diversity of exposure is an important component of a good learning environment, one that parents and teachers should attempt to incorporate. Yet it remains unclear how the infant recognizes that exemplars spoken by different talkers should ever be mapped onto the same word (that is, how exemplars from multiple talkers saying *cat* get treated as different tokens of the same word whereas the same talker saying both *cat* and *pat* does not).

The importance of variability for generalizing across talker voice and affect raises questions regarding its importance for other areas of word learning as well. If an infant hears frequent mispronunciations, will that infant begin to ignore phonetic details about the word (e.g., accepting *tat* as a variant of *cat*, because a prior speaker had confused the two)? In theory, variability could also drive infants' developing ability to perceive speech in noise: Perhaps experiencing different types of noise allows infants to learn that noise itself is not part of a word's representation. Neither of these possibilities has been explored experimentally, but examining individual differences among infants might provide a means of investigating such issues.

Another question is whether this developmental change occurs globally (on the processes that act upon representations) or on a word-by-word basis (on the representations themselves). As infants gain greater linguistic sophistication, they may begin processing lexical information differently across the board. Alternatively, as an infant gains more experience with a particular word, the representation for that word alone might change, allowing for better recognition on the basis of partial or noisy information. Evaluating the nature of developmental change in word recognition could inform the controversy between representation and process more generally.

Current research has focused on demonstrating that infants' representations contain a great deal of acoustic detail and suggests that variability in the input may be a critical factor in enhancing children's generalization. Future research needs to explore the mechanisms by which this representational change occurs.

Recommended Reading

Jusczyk, P.W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press. A clearly written and relatively comprehensive review for readers who wish to expand their knowledge on how infant speech perception leads to language acquisition.

Werker, J.F., & Yeung, H.H. (2005). Infant speech perception bootstraps word learning. *Trends in Cognitive Science*, *9*, 519–527. A more recent review of current theory on the relationship between speech perception and infant word learning.

Acknowledgments—I thank David Rossell, the editor, and two anonymous reviewers for helpful comments on earlier drafts of this manuscript. This work was supported in part by the National Science Foundation under grant number BCS 0642294.

REFERENCES

Berwick, R.C. (1985). *The acquisition of syntactic knowledge*. Cambridge, MA: MIT Press.

Braine, M.D.S., & Brooks, P.J. (1995). Verb argument structure and the problem of avoiding an overgeneral grammar. In M. Tomasello & W.E. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs*. Hillsdale, NJ: Erlbaum.

Fennell, C.T., & Werker, J.F. (2004). Infant attention to phonetic detail: Knowledge and familiarity effects. In A. Brugos, L. Micciulla, & C.E. Smith (Eds.), *Proceedings of the 28th Annual Boston University Conference on Language Development* (pp. 165–176). Somerville, MA: Cascadilla Press.

Fernald, A., Swingley, D., & Pinto, J.P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, *72*, 1003–1015.

Gómez, R.L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*, 431–436.

Hollich, G. (2006). Combining techniques to reveal emergent effects in infants' segmentation, word learning, and grammar. *Language & Speech*, *49*, 3–19.

Hollich, G., Golinkoff, R.M., & Hirsh-Pasek, K. (2007). Young children associate novel words with complex objects rather than salient parts. *Developmental Psychology*, *43*, 1051–1061.

Houston, D.M., & Jusczyk, P.W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception & Performance*, *26*, 1570–1582.

Jusczyk, P.W., & Aslin, R.N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*, 1–23.

Jusczyk, P.W., Pisoni, D.B., & Mullennix, J. (1992). Some consequences of talker variability on speech perception by 2-month-old infants. *Cognition*, *43*, 253–291.

Naigles, L.R., & Hoff, E. (2006). Verbs at the very beginning: Parallels between comprehension and input. In K. Hirsh-Pasek & R.M. Golinkoff (Eds.), *Action meets word: How children learn verbs*. Oxford, UK: Oxford University Press.

Newman, R.S. (2006). Perceptual restoration in toddlers. *Perception & Psychophysics*, *68*, 625–642.

Newman, R.S., & Jusczyk, P.W. (1996). The cocktail party effect in infants. *Perception & Psychophysics*, *58*, 1145–1156.

Singh, L., Bortfeld, H., & Morgan, J.L. (2002, November). *Effects of variability on infant word recognition*. Paper presented at the 26th Annual Boston University Conference on Language Development, Somerville, MA.

Singh, L., Morgan, J.L., & White, K.S. (2004). Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language*, *51*, 173–189.

Stager, C.L., & Werker, J.F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*, 381–382.

Swingley, D., & Aslin, R.N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*, 480–484.

Werker, J.F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning & Development*, *1*, 197–234.

Yip, M. (2002). *Tone*. Cambridge, UK: Cambridge University Press.