Journal of Memory and Language 88 (2016) 117–132

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Modeling the role of distributional information in children's use of phonemic contrasts

Erik D. Thiessen [a,*], Philip I. Pavlik Jr. [b]

[a] Carnegie Mellon University, United States
[b] University of Memphis, United States

ARTICLE INFO

Article history:
Received 1 December 2014
revision received 13 January 2016
Available online 18 February 2016

Keywords:
Statistical learning
Word learning
Exemplar memory model
Memory
Abstraction
Language

ABSTRACT

Between the first and the second year of life, children improve in their ability to use phonemic contrasts when learning label–object pairings. This improvement may be related to children's experience with the distribution of phonemes across lexical forms. Because phonemes typically occur in different lexical frames (e.g., /d/ and /t/ in "doggy" and "teddy" rather than "doggy" and "toggy"), familiarity with words makes similar phonemes more distinct through acquired distinctiveness. In a series of simulations, we demonstrate that English input has the distributional characteristics necessary to facilitate use of phonemic contrasts as a function of increasing familiarity with the lexicon. Further, these simulations support a novel prediction: that less common phonemes should take longer to be used productively. We tested this prediction with children between 18 and 25 months, and found that the relatively infrequent /s/ and /z/ contrast takes longer to emerge than frequent contrasts such as /b/–/d/ or /d/–/t/.

By the first birthday, the average monolingual infant is already familiar with the meaning of nearly 100 words (Fenson et al., 2002). But despite infants' success in discovering word–object associations at this age, they appear to have difficulty taking advantage of phonemic distinctions when making connections between word forms and meaning. One striking example of this is infants' failure to respond differentially to lexical forms that differ by only a single phoneme (i.e., *minimal pairs*) in label–object association tasks. For example, 14-month-old infants who have learned (via habituation) that a novel object is associated with the label /da/ respond equivalently when that object is labeled /ta/ as when it is labeled /da/. This failure to differentiate between the labels is not limited to /da/ and /ta/, and has been replicated with a variety of phonemic contrasts in word–object association tasks (e.g., Pater, Stager, & Werker, 2004; Thiessen & Yee, 2010). Infants' failure to respond to these phonemic differences is not due to an inability to hear them, but is instead specific to settings where they are asked to use these contrasts to differentiate word meanings (e.g., Stager & Werker, 1997; Swingley & Aslin, 2000; Thiessen, 2007; Yoshida, Fennell, Swingley, & Werker, 2009). Instead of a perceptual failure, infants' failures to use phonemic differences in label–object association tasks appear to results from a difficulty in *making use* of some phonemic distinctions that they can perceive and encode (e.g., Shvachkin, 1973; Thiessen, 2011).

That is, though infants perceive and encode the phonemic distinctions that are relevant to their native language when they are building a lexicon, they initially fail to treat these phonemic distinctions as signifying the distinction between tokens of different lexical categories (Swingley & Aslin, 2007). Over the course of development, this difficulty is substantially alleviated. By 18 months, infants use phonemic distinctions (at least, the word-initial stop consonant contrasts typically used in laboratory tasks)

* Corresponding author at: Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213, United States.
E-mail address: thiessen@andrew.cmu.edu (E.D. Thiessen).

http://dx.doi.org/10.1016/j.jml.2016.01.003
0749-596X/

contrastively in word–object association tasks (e.g., Thiessen, 2007; Thiessen & Yee, 2010). Three potential explanations have been forwarded for this developmental change. The first is a capacity account. On this account, young infants' failures are due to the fact that the task is demanding – it requires infants to encode the label, the object, and the link between them. Because young infants have less capacity than older infants and adults, they cannot encode all of this information simultaneously, and subsequently fail to detect subtle changes in the identity of the label. Older infants succeed because they have more capacity (Werker, Fennell, Corcoran, & Stager, 2002). The second is a social inference account. From this perspective, the reason that infants fail is that the label–object tasks used in the lab lack sufficient social support for infants to react to the labels as though they are meaningful. As such, infants ignore distinctions that they would treat as informative in a "real" or "linguistic" label–object association setting (Fennell & Waxman, 2010). While it is clear that social factors influence infants' word learning, this perspective does not make strong claims about what differentiates younger infants (who fail to use phonemic distinctions in laboratory tasks) from older infants (who succeed).

A final account suggests that infants' success and failure in using phonemic distinctions in laboratory tasks is due to their experience with the distributions of phonemes in their native language. In particular, Thiessen and colleagues (Thiessen, 2007, 2011; Thiessen & Pavlik, 2013; Thiessen & Yee, 2010) have proposed that experience with lexical forms helps to differentiate phonemic contrasts. On this account, infants are encoding a great deal of acoustic information associated with lexical forms. In addition to phonemic identity, for example, infants also encode indexical information about the speaker of a word (e.g., Houston & Jusczyk, 2003; Singh, 2008; Werker & Curtin, 2005). Initially, it may not be apparent which aspects of the acoustic variability in the encoded word forms are relevant for differentiating among different tokens of spoken words. Experience with lexical forms is informative to the extent that these forms help infants resolve the ambiguity inherent in the perceptual input and lead infants to weight phonemic distinctions more heavily (e.g., Swingley, 2009). This resolution occurs because experience with lexical forms is not random. Instead, infants are especially likely to encounter phonemes in distinct lexical contexts (such as /d/ and /t/ in *doggy* and *teddy*) as they develop a lexicon. Compared with the adult lexicon, children's lexicons contain fewer words where phonemes occur in identical contexts (i.e., fewer minimal pairs, such as /d/ and /t/ in *dip* and *tip*). For example, Swingley and Aslin (2007) found that over two thirds of the words in the vocabularies of 18-month-old Dutch learning infants had no minimal pair neighbors. Similarly, there are no single-feature minimal pairs in the first 50 words that children are most likely to comprehend (Caselli et al., 1995).

The experience of phonemic contrasts in distinct lexical contexts may serve to differentiate the contrasts due to a process known as *acquired distinctiveness*. Two similar stimuli, when paired with distinct outcomes, become more differentiable (e.g., Hall, 1991). That is, if an organism has difficulty differentiating between stimuli *A* and *B* (i.e., two similar phonemes), they can be repeatedly paired with two more easily distinguished outcomes *X* and *Y* (e.g., *X* might be reward and *Y* punishment), such that the organism consistently experiences *AX* and *BY* pairings. Over time, these pairings reinforce the original (difficult) distinction between *A* and *B* and the distinction becomes more robust. In the child's developing lexicon, phonemes that are initially difficult to distinguish become more differentiable as they are paired with distinct lexical contexts. Because children know so few minimal pair words, they are unlikely to experience two phonemes in identical contexts (like /d/ and /t/ in *dip* and *tip*). Instead, they experience phonemes primarily in distinct contexts, and this experience with lexical could potentially help to differentiate similar phonemes.

To test this hypothesis, Thiessen (2007) conducted a laboratory training procedure intended to facilitate children's use of a phonemic contrast in a word–object association task. In that procedure, 15-month-olds who typically fail to respond differentially to the /d/–/t/ distinction in a word–object association task were exposed to the contrast in distinct lexical contexts, /dabo/ and /tagu/. After exposure to these labels, infants succeeded in responding differentially to the labels /da/ and /ta/ (for a replication, see Thiessen, 2011). This result is not simply due to increased familiarity with /d/ and /t/. When infants were exposed to these consonants for the same amount of time in *identical* lexical contexts (such as /dagu/ and /tagu/), they showed no benefit from the exposure and continued to respond to /d/ and /t/ as though they were interchangeable. These results are consistent with the hypothesis that exposure to phonemes in distinct lexical contexts helps to make similar phonemes more differentiable. More generally, this suggests that the distribution of phonemes in lexical forms, and infants' increasing familiarity with those lexical forms, plays an important role in infants' improving ability to make use of phonemic contrasts in word–object association tasks (e.g., Stager & Werker, 1997; Werker et al., 2002).

Of the three accounts discussed here, the distributional experience account makes a unique prediction: it suggests that children's ability to make use of particular phonemic contrasts in a label–object association task should be predictable from the frequency and distribution with which children have experienced those specific contrasts. That is, when two phonemes are similar to each other – for example, phonemes that differ by only a single phonetic feature such as voicing, which we will refer to as "minimal pair" phonemes – children need to experience the phoneme in distinct contexts for the members of the phoneme to become distinct enough to respond to them differentially in a label–object association task. This will take different amounts of time for different phonemes, as a function of the frequency with which children experience them in their language, such that children should succeed with more frequent contrasts (such as /d/ and /t/) before they succeed with less frequent contrast (such as /s/ and /z/). By contrast, the capacity account (Werker et al., 2002) suggests that children should succeed with all phonemes once they have enough capacity to encode the