THE ROBUSTNESS OF INFANTS' EARLY WORD REPRESENTATIONS

Marieke van Heugten and Elizabeth K. Johnson

Dept. of Psychology, University of Toronto, 3359 Mississauga Road N., Mississauga, ON, L5L 1C6, Canada
marieke.vanheugten@utoronto.ca

1. INTRODUCTION

The speech stream contains no one-to-one mapping between acoustic form and lexical entry. Indeed, factors such as idiolect and speaking rate ensure that no word is ever produced in an identical fashion twice. Nevertheless, listeners effortlessly understand spoken language. In this paper, we explore how infants, who have yet to develop a mental lexicon, start recognizing acoustically distinct realizations of words.

Although infants can cope with some natural speech variation (Jusczyk & Aslin, 1995), recent research suggests that word recognition abilities are initially very limited (Houston & Jusczyk, 2000; Singh, Morgan, & White, 2004). At 7.5 months of age, infants familiarized with *bike* and *pear*, for example, listened longer to passages containing these words as opposed to passages containing novel words (e.g., *hat* and *tree*), indicating that they had recognized these words when embedded in context. This effect, however, only holds for words that are acoustically similar during familiarization and test. When the voice changes gender from a male to female speaker or when the affect changes from a happy to a neutral tone, they no longer appear to map the word forms onto the same underlying representations. 10.5-months-olds, in contrast, do generalize across distinct word tokens, indicating that this abstraction ability develops over time. Exemplar-based models have taken this developmental pattern from the early lack of perceptual constancy towards more adult-like word recognition to indicate that words may first be learned in an instance-specific fashion, with generalizations only occurring once a sufficient number of word tokens have been encountered.

In this study, we examine the possibility that past work has underestimated infants' early word recognition abilities. We propose that presenting infants with brief exposure to disembodied unfamiliar voices producing isolated words is not an ecologically valid measure of infants' capabilities. Even adult listeners have difficulty when tested under similar circumstances (Nygaard, Sommers, & Pisoni, 1994). In order to examine this possibility, two groups of infants were tested. The first group was familiarized with passages recorded by their mother. They were then tested on lists of isolated word tokens spoken by their father. The second group of infants heard the same passages and test lists spoken by an unfamiliar speaker (i.e. another infant's parents). Including both groups allows us to test for the possibility that infants can cope with variability in the realization of words more effectively when hearing familiar voices. Following previous studies (e.g., Houston &

Jusczyk, 2000), longer listening times for lists with the familiarized as opposed novel words would indicate cross-gender word recognition. We thus predicted that infants presented with familiar voices would exhibit such pattern and succeed in recognizing words despite the change in gender between familiarization and test. Infants presented with unfamiliar voices, in contrast, should perform in a fashion comparable to past studies and fail to recognize words across different-gender speakers.

2. METHOD

2.1 Participants

Fourty-eight normally developing monolingual English-learning 7.5-month-olds were tested in this study (age range: 220 - 248 days; 25 girls). An additional eleven infants were tested, but excluded from the analysis due to fussiness.

2.2 Materials

Prior to test, twenty-four mothers and fathers whose infants were recruited to participate in the study were audio-taped. Mothers recorded two six-sentence passages in infant-directed speech. Each passage contained a target word (boat, cup, pear, or toque) that occurred once every sentence (see Table 1). For each mother, a familiarization video was created consisting of a picture of the mother zooming in and out combined with three alternating repetitions of each of the two recorded passages. All familiarization videos contained 18 tokens of each of the target words. Fathers recorded test lists with 15 repetitions of the isolated target.

2.3 Procedure

Infants were seated on their parent's lap in front of a TV screen in a sound-attenuated booth. After viewing the familiarization video, the test phase started. An adapted version of the Headturn Preference Procedure (as used by Jusczyk & Aslin, 1995) was used to test infants' recognition of the familiarized words across different genders. First, a red light in front of the infant started flashing. Once the infant oriented towards this light, one of the two lights at the side panels started to flash. As soon as the child turned towards this flashing light, the word list started playing from the loudspeaker mounted underneath the light. Trials were either played until the end of the list or until the infant looked away for two seconds. The computer recorded infants' orientation times to each of the trials.

Table 1. Example passage from familiarization phase

Her boat had white sails. This girl will steer my big boat. That horn on the boat was really loud. He bought himself a new red boat. His boat could go quite fast. We always store your boat in our garage.

2.4 Design

To test the effect of word familiarity, half of the infants were presented with their own parents' voices, while the other half were yoked pairs listening to unfamiliar voices. The target words used during familiarization (*boat* and *tuque*, *boat* and *cup*, *tuque* and *pear*, or *pear* and *cup*) were counterbalanced across conditions. The four test lists were identical across conditions and were presented three times each in quasi-random order.

3. RESULTS

Mean listening times to familiarized and novel words were calculated for each infant separately[1]. Across all subjects tested on their parents' voices, the average listening time to lists with familiarized words was 10.18 s (SEM = .57) and to lists with novel words 9.45 s (SEM = .61). For subjects tested on unfamiliar voices, the average listening time to lists with familiarized words was 10.31 s (SEM = .46) and to lists with novel words 9.51 (SEM = .53; see Figure 1). A mixed 2x2 ANOVA with Word Familiarity (familiarized vs novel words) as within-subject and Voice Familiarity (own parents' voices vs unknown voices) as between-subject factor revealed a main effect of Word Familiarity ($F(1,46) = 4.224$; $p = .046$), indicating that infants listened longer to word lists containing familiarized as opposed to novel words. No other significant main effects or interactions were found (all Fs < 1).

4. DISCUSSION

Contrary to our predictions, 7.5-month-olds recognized familiarized words across speakers of different genders, regardless of whether they were presented with familiar or unfamiliar voices. Infants thus seem to generalize across acoustically distinct instances of the same word, suggesting that their early speech perception abilities are more advanced than generally assumed. While this finding is surprising, differences between past and present studies may explain the current findings. Unlike all past studies showing limitations to infants' early generalization capacities, infants in this study were familiarized with passages rather than words, increasing the overall exposure to talker-specific
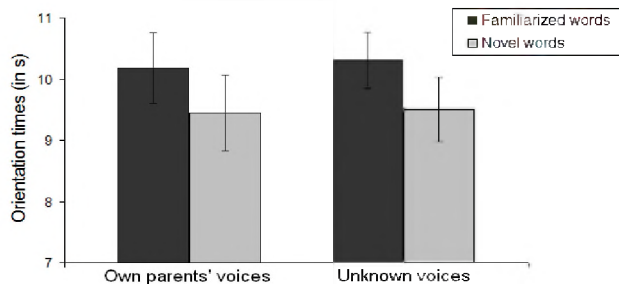

Figure 1. Mean orientation times to familiarized and novel words (and SEMs) broken down by Voice Familiarity

---

[1] Orientation times more than 2.5 standard deviations away from the infant's mean (3 out of 576 data points) were discarded.

idiosyncrasies. This type of context could be of great help when extracting the intended linguistic form from the speech signal. Contextual information, in other words, may have enhanced speaker adaptation, thereby facilitating word encoding. Support for this view comes from a study with French 8-month-olds, showing that newly presented words are learned better when infants are familiarized with passages than with isolated word tokens (Polka et al., 2008). Another difference between this and earlier studies involves the materials. First, parents' rather than actors' voices were used in this experiment. Second, testing infants on couples' voices may have facilitated word recognition. Vocal convergence might have led to greater acoustic similarity between two partners' than between two randomly selected voices, which in turn may have assisted infants' generalization across the two voices.

Although infants' word recognition in this study was not facilitated by the use of familiar voices, it is very well possible that their robust generalization abilities masked any such effect. In follow-up studies, we are therefore examining the potential advantage of familiar over unfamiliar voices under more challenging listening conditions. Nonetheless, the finding that infants do recognize words across acoustically distinct realizations is intriguing. Models of early speech perception have considered infants' initial lack of generalization as evidence for an episodic memory for speech. When familiarized with more naturalistic stimuli (i.e. passages rather than words), however, infants handle variability in the speech signal more competently than previously thought. Future models of word recognition thus need to take into account infants' abilities under greater acoustical validity.

REFERENCES

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, 28*, 1-23.

Nygaard, L.C., Sommers, M.S., & Pisoni, D.B. (1994). Speech perception as a talker-contingent process. *Psychological Science, 5*, 42-46.

Polka, L., Proulx, J., Mersad, K., Iakimova, G., Sundara, M. & Nazzi, T. (2008). *Speech segmentation in French-learning infants: Language-specific and dialect-specific patterns.* Talk presented at the 16th International Conference on Infant Studies, Vancouver, Canada.

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.