INFANTS' SPEECH SEGMENTATION

Student Authors

Authors (left to right) **Meredith Smith, Allison Kish, Katie Courtaney, Serina Thottichira, Ambry Roberson, Tori Gallo, Amy Jarvis, Lauren King,** and **Adi Dilger** are all undergraduate students in the Speech, Language, and Hearing Sciences Department at Purdue University. All of the authors have worked in the Purdue Infant Lab under the supervision of Dr. Amanda Seidl. This study started as a research project undertaken while completing an independent study course associated with the Purdue Infant Lab.

Mentors

Amanda Seidl is an associate professor in the Speech, Language, and Hearing Sciences Department. The overarching goal of her research is to discover how language comes to the child. Seidl's current research program focuses on early predictors of language. Specifically, this work explores the ways in which measures of early speech perception, production, and the input to the child relate to later language in both typical development and in children at risk for autism spectrum disorders.

Rana Abu-Zhaya is a PhD student in the Speech, Language, and Hearing Sciences Department. She has worked as a speech-language pathologist in Israel and is exploring the role of dyadic interactions in the course of language acquisition for her PhD work, focusing on effects of caregiver touch on infant word segmentation and learning.

Abstract

Various measures of infant responsiveness have been shown to predict child outcomes. Despite this extensive research, there is no work examining links between infant responsiveness during caregiver-infant interactions with infants' ability to perform basic linguistic tasks. One key task in early linguistic development is word segmentation, an achievement that allows infants to build their mental dictionaries. We hypothesized that infants' responsiveness to caregiver facial expressions might be related to their word segmentation ability.

In order to test this hypothesis, mothers came into the lab and were videotaped reading books containing target words to their 5-month-old children. After the infants were read to, we tested their listening preference for words in the books, as well as novel words; this test yielded a preference score (preference for familiar vs. unfamiliar words). We also used the videotaped reading to code facial expressions for both infant and caregiver, and subsequently, we tabulated occasions where synchronous facial expressions occurred for each member of the dyad.

We then examined possible correlations between our preference score and measures gleaned from the dyadic facial expression coding. Although neither the number of infant-led synchronous facial expressions nor the total number of facial expressions produced by either member was significantly correlated with preference score, our measure of synchronous facial expressions led by the caregiver was highly correlated with preference score. Thus, results support the hypothesis that infant responsiveness during caregiver-infant interaction, as indexed by synchronous facial expressions with caregivers, may be related to language learning ability.

Kish, A., Courtaney, K., Dilger, A., Gallo, T., Jarvis, A., King, L., . . . Thottichira, S. (2014). Infants' speech segmentation: The impact of mother-infant facial synchrony. *Journal of Purdue Undergraduate Research, 4,* 10–15. http://dx.doi.org/10.5703/jpur.04.1.02

Keywords

speech perception, facial expressions, infant-caregiver interaction, facial synchrony, word segmentation, touch, development, reciprocity

http://dx.doi.org/10.5703/jpur.04.1.02

INFANTS' SPEECH SEGMENTATION:

The Impact of Mother-Infant Facial Synchrony

Allison Kish, Katie Courtaney, Adi Dilger, Tori Gallo, Amy Jarvis, Lauren King,
Ambry Roberson, Meredith Smith, and Serina Thottichira, Speech, Language, and Hearing Sciences

INTRODUCTION

Word segmentation, or the ability to find words in the continuous flow of speech, has been shown to be an important element in early linguistic development. Without skills in word segmentation, infants cannot begin to build a lexicon (or mental dictionary). Given that the words spoken to infants primarily consist of continuous speech (Johnson, Seidl, & Tyler, 2014; van de Weijer, 1998), all infants must learn how to segment the speech stream in a language-specific manner. (Anyone who has ever listened to a completely unfamiliar language and attempted to locate individual words in that language can agree that this is a difficult task.) Further, there is individual variability in infants' word segmentation ability between 7 and 12 months of age, such that infants' word segmentation ability within this age range predicts later toddler and preschooler language ability (Junge, Kooijman, Hagoort, & Cutler, 2012; Newman, Bernstein Ratner, Jusczyk, Jusczyk, & Dow, 2006; Singh, Reznick, & Xuehua, 2012).

We sought to explore the source of this individual variability in word segmentation ability and hypothesized that it may be related to infant attentiveness and responsivity. Infant responsiveness to caregivers is present from birth (Harrist & Waugh, 2001) and is often instantiated by an infant's responsiveness to her caregiver's facial expressions. In this study, we explored whether infants' attention and responsivity, as indexed by their responsiveness to their caregivers' facial expressions, might be related to their ability to segment the speech stream produced by their caregivers. This is an interesting question, since while various measures of infant responsiveness have been shown to predict later cognitive outcomes in children, no research has been conducted linking infant responsiveness to caregiver-infant interactions with infants' ability to perform basic language tasks such as word segmentation.

As mentioned, our research question addresses how infants' attention to caregivers' facial expressions may impact infants' ability to segment their caregivers' speech. In order to explore this question, 12 mothers of 5-month-old infants were invited to come into our lab with their children. We chose to examine 5-month-old infants because this is an age before which word segmentation is typically attained; thus, we would expect a large range of individual variability at this age.

Mothers read two books focused on different categories—animals and body parts—to their infants. We chose to examine these two categories since (a) these are common topics in infant books and (b) we expected that caregivers might be more expressive in discussing one category of objects with their infants than another category of objects. We hypothesized that this expressiveness might impact infants' responsivity and hence their ability to find words in the speech stream. Caregiver-infant interactions during book readings were videotaped in the lab for later coding. Note that during the book reading, infants were familiarized with target words such as "chin" or "cat," embedded in sentential contexts in the books, such that all words in the books would have to be segmented.

Following the book reading, each mother-infant dyad was moved to a room where we tested the infants' ability to recognize words from the books. Specifically, this procedure tested the infants' preference for the target words they were familiarized with during the book reading (e.g., “chin” and “cat”), compared with their preference for novel or unfamiliar words, which were not present within the books (e.g., “knee” and “dog”). A single preference score was yielded from this task that determined infants’ ability to segment the familiar target words.

Videotaped caregiver-infant interactions during the book reading task were coded to tabulate instances of infant-caregiver synchrony in facial expressions. Coded facial expressions included: eyebrow raising, smiling, grimacing, and lip rounding. We chose to explore synchrony in these expressions since, according to previous research (Harrist & Waugh, 2001), the ability to achieve synchrony “may represent a crucial developmental achievement for significant dyadic relationships, one that facilitates social, emotional, and cognitive growth for the child” (p. 555). Lastly, correlations between our listening preference score and the measures collected from our dyadic facial expression coding were examined. If our hypothesis is correct, then our research could aid in the understanding of how individual differences in infant word segmentation (and subsequent vocabulary growth) relates to infant attentiveness and responsiveness during caregiver-infant interactions.

METHODS

Participants

Twelve mother-infant dyads participated in the study when their monolingual, typically developing, English-learning infants were about 5 to 6 months of age (mean = 5.26 months; range = 4.34 months to 5.79 months). Participants were recruited from flyers posted in the Greater Lafayette region as well as through the birth announcements in the newspaper. All mothers gave informed consent and all infants received a book or toy for their participation in the study.

Familiarization Stimuli

For counterbalancing purposes, eight books were constructed by the researchers for caregivers to read to their infants. Four of these books were on body parts and the other four were on animals (see Table 1 for an illustration of the words in four of the books of each type). Each book included four target words. Each caregiver was given one of each type of book (body part, animal), thus familiarizing each infant with a total of eight words. These books contained target words such as “chin” (body part) or “cat” (animal) that were embedded in the sentential contexts. To ensure uniformity, the books were constructed with the same sentential structures and photographic layouts. They varied only in their target words (see Figures 1 and 2). Each book included ten pages with typed text in which target words were located at the end of the sentences. In addition, pictures in the books were carefully chosen from photographic databases on the web in order to control for extraneous variables such as salience (Google Images; www.pictures.com; *National Geographic*; PeopleImages).

Body Part Books			
B1	B2	B3	B4
Belly (SW)	Tummy (SW)	Finger (SW)	Eyebrow (SW)
Nose (S)	Eye (S)	Mouth (S)	Ear (S)
Chin (S)	Waist (S)	Knee (S)	Heel (S)
Leg (S)	Feet (S)	Toe (S)	Hand (S)

Animal Books			
A1	A2	A3	A4
Camel (SW)	Puppy (SW)	Lion (SW)	Hippo (SW)
Bear (S)	Bird (S)	Frog (S)	Snake (S)
Cat (S)	Horse (S)	Mouse (S)	Dog (S)
Sheep (S)	Cow (S)	Duck (S)	Pig (S)

Table 1. Animal and body part books with stress pattern (S = strong, SW = strong-weak).

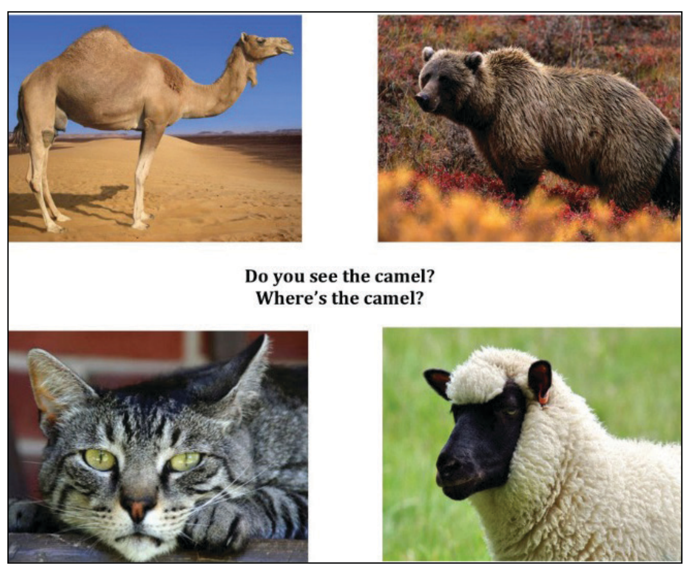

Figure 1. Example page from animal book.

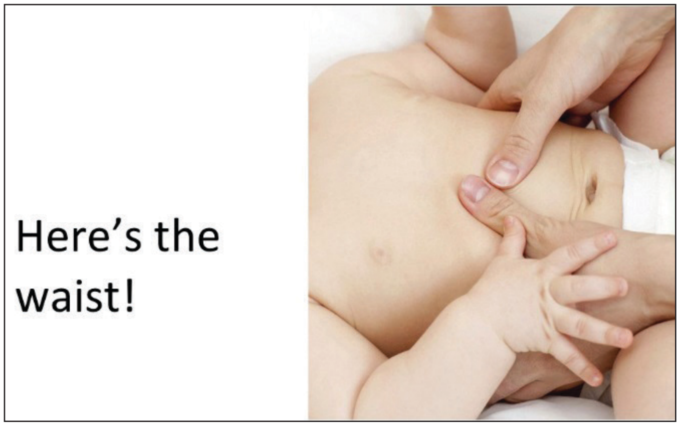

Figure 2. Example page from body part book.

Test Stimuli

Test stimuli consisted of recordings of the target words presented in isolation ("chin" bracketed on both sides by pauses) repeatedly, as well as words not presented in the books recorded in the same manner (e.g., "knee"). These words were recorded in infant-directed speech by a young female native English speaker and were presented at a comfortable volume (72 decibels).

Familiarization Procedure

Caregivers were instructed to read the two books to their infants two times each. During the book reading, mothers sat in a chair opposite their infants, who were seated in a high chair during the interaction (see Figure 3). These mother-infant interactions were videotaped for later coding using two cameras, one focused more on the infant and the other focused more on the mother. Mothers wore a clip-on microphone that was wirelessly connected to one of the video recorders.

Figure 3. Caregiver-infant interaction coding.

After familiarization was complete, we used ELAN software (Brugman & Russel, 2004) to code the videos for caregiver-infant facial expression synchrony. Specifically, one group of students coded infants' facial expressions while another group of students coded mothers' facial expressions. These expressions included eyebrow raising, smiling, grimacing, lip rounding, and mouth opening. Next, a third group of coders tabulated the total number of facial expressions for mothers (M-total) and infants (I-total) as well as occasions where expressions were synchronous and initiated by the mother (MI-total) and synchronous and initiated by the infant (II-total). Synchrony here was defined as the same facial expressions (e.g., smiles for both mother and infant) occurring for each member of the dyad within a 1-second window.

Testing Procedure

After the infants were familiarized with the target words through the readings, we brought them into a testing room and used a procedure known as the Headturn Preference procedure (HPP; see Figure 4). Here we tested infants' listening preferences for words that they heard embedded in the books (e.g., "cat" and "nose" for infants familiarized with A1 and B1 in Table 1), as well as novel words that did not appear in the books (e.g., "knee" and "mouse" for infants familiarized with books A1 and B1 in Table 1).

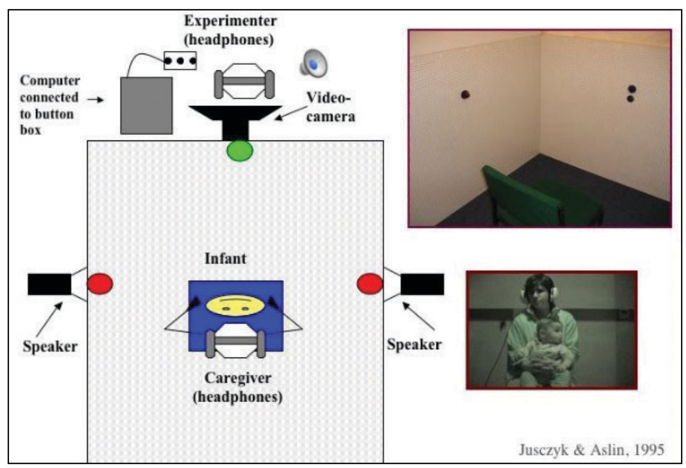

Figure 4. Three perspectives on the HPP.

The infant sat on his or her caregiver's lap for the HPP. The infant and caregiver were located in a room that contained lights positioned on the front and two sidewalls. The lights on the sidewalls had speakers directly behind them. Caregivers and experimenters wore Peltor aviation headphones and listened to a combination of loud masking music and white noise so as not to influence the infants' behaviors. To begin each trial, the experimenter attracted the baby's attention by flashing a green light located in front of the infant. Once the infant focused his/her attention on the front light, one of the two red lights located on either side of the infant flashed. When the infant looked at the illuminated light, the experimenter signaled the computer to play a sound from the light's respective speaker. The amount of time the infant looked at the side from which the speaker played the familiar or unfamiliar words was calculated and recorded. When the infant oriented away from the light for more than 2 seconds, the sounds and light were extinguished and the next trial began. There were a total of 16 test trials of two types (8 familiar, 8 unfamiliar) presented in 1 block, with trial order and light side fully randomized by a computer program. The computer recorded our dependent measure: total looking time to the familiar words minus the infants' looking time to the unfamiliar words. This was calculated overall for animal and body part words separately.

RESULTS

We ran a regression to explore the relationship between infants' preference scores in the HPP task (the difference score computed by calculating each infant's attention to familiar words minus his or her attention to unfamiliar words) and total number of facial expressions for mother (M-total) and infant (I-total), as well as reciprocal facial expressions led by the mother (MI-total) or led by the infant (II-total). This regression revealed that, while neither total number of facial expressions from the mother or the infant (*p*s > .17) nor infant-initiated total (II-total) predicted infants' word segmentation ability, MI-total was marginally related to infants' word segmentation ability ($p = .07$).

Given this marginal effect, and the fact that we had two books with different words from different categories (animals and body parts) that we predicted may cause caregivers to use different social cues in interacting with their infants, we ran a series of correlations to examine whether caregivers' facial expressions were related to infants' preference scores for each body part word and animal word separately. While neither the number of facial expressions for mother and infant nor the number of reciprocal facial expressions initiated by the infant or the mother predicted infant preference scores for animal words (all *p*s > .43), mother-initiated facial expressions were related to infants' difference scores for body part words ($r = .69$, $p < .01$; all other effects for body part words were not significant, all *p*s > .08). This significant correlation is shown in Figure 5.

DISCUSSION

Results of this study suggest that infants who are more responsive to their caregivers' facial expressions are better at finding body part words in the continuous stream of speech. Interestingly, infants' ability to find words in the continuous stream of speech does not

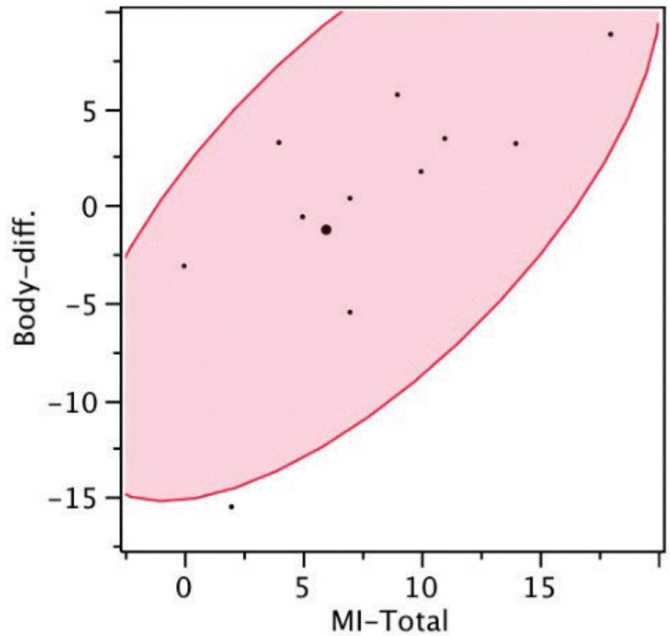

Figure 5. The correlation between an infant's preference for familiar body part words and responsiveness to their mother's facial expressions.

appear to be related to the overall number of facial expressions by mother or infant, nor does it appear to be related to caregivers' responsiveness to their infants' facial expressions, but only the infants' responsiveness to their caregiver.

It is worthy of note that the mimicking facial expressions described here is not an ability exclusive to humans, but an ability that is common to other mammals (Campbell & de Waal, 2014). Therefore, facial expression reciprocity is likely biologically programmed in mammals, and the amount of mimicking, in regards to facial expressions, often indicates the degree of familiarity in human interactions. However, the question that we explore in this paper (whether facial expression reciprocity is linking with linguistic ability) is necessarily unique to humans since language is unique to the human species. Thus, our research suggesting that the mimicking of caregiver facial expressions also indicating an attention to caregivers' words is a novel contribution to this literature.

Interestingly, our data suggest that the degree of mother-initiated reciprocity is most closely related to the segmentation of body part words specifically. Why did we find that facial expression reciprocity initiated by the mother and responded to by the child predicted infant ability to find body part words in continuous speech, but not animal words? One hypothesis might be that caregivers produce more facial expressions when reading the body part books; however, if this hypothesis were correct, an overall correlation would have been found between total number of caregiver facial expressions and infant segmentation. We found no such effects. A second hypothesis might be that infants' responsiveness and increased segmentation to body part words might be due to caregivers interacting with infants in a more direct manner when reading the body part books than they did when reading the animal books. Specifically, we suggest that it may be that caregivers provided additional tactile cues when reading the body part books, and these cues may have served to boost the impact of reciprocity for infants.

While our results are suggestive, further research regarding word segmentation and tactile cues would strengthen the validity of the latter hypothesis. Other improvements or suggestions concerning our study include recruiting a larger sample size and implementing the use of an automatic coding program for facial expressions of both the infant and the caregiver. A larger sample size also could strengthen validity, as well as increase the diversity of the population in our sample. Moreover, an automatic coding program could yield more accurate temporal coding records, and this could be followed by manual review of the coded videos.

Lastly, our research could be applied to the individuals at risk for autism spectrum disorders (e.g., infant siblings of children with known ASD diagnoses) in order to pinpoint early indicators of autism in comparison to typically developing children. Specifically, infants who later go on to be diagnosed with autism spectrum disorders often have difficulty with social cues and facial expression reciprocity (Apicella et al., 2013). Our results showed a relationship between mother-initiated facial expression synchrony and body part word segmentation abilities in typically developing infants. Given that infants who go on to develop autism spectrum disorders often are not able to attend to the social and facial cues given by their mothers, the results of this study lead us to hypothesize that this difficulty in social reciprocity might contribute to the delays in their development of important lexical information. Specifically, if body part words provide a key toehold for lexical acquisition and social cues aid significantly in their acquisition, then infants who fail to attend to these social cues may have a deficit in lexical acquisition that begins with the acquisition of these body part words between 5 and 6 months of age. Only future work will allow us to address these interesting hypotheses.

ACKNOWLEDGMENTS

The authors would like to thank all the families who participated in our study.

REFERENCES

Apicella, F., Chericoni, N., Costanzo, V., Baldini, S., Billeci, L., Cohen, D., & Muratori, F. (2013). Reciprocity in interaction: A window on the first year of life in autism. *Autism Research and Treatment, 2013,* 1–12. http://dx.doi.org/10.1155/2013/705895

Brugman, H., & Russel, A. (2004). Annotating multimedia/multi-modal resources with ELAN. *Proceedings of LREC 2004, Fourth International Conference on Language Resources and Evaluation* (LREC 2004), pp. 2065–2068.

Campbell, M. W., & de Waal, F. B. M. (2014). Chimpanzees empathize with group mates and humans, but not with baboons or unfamiliar chimpanzees. *Proceedings of the Royal Society B: Biological Sciences, 281*(1782). http://dx.doi.org/10.1098/rspb.2014.0013

Harrist, A. W., & Waugh, R. M. (2001). Dyadic synchrony: Its structure and function in children's development. *Department of Human Development and Family Science*, *22*(4), 555–592. http://dx.doi.org/10.1016/S0273-2297(02)00500-2

Johnson, E. K., Seidl, A., Tyler, M. D. (2014). The edge factor in early word segmentation: Utterance-level prosody enables word form extraction by 6-month-olds. *PLOS ONE, 9*(1). http://dx.doi.org/10.1371/journal.pone.0083546

Junge, C., Kooijman, V., Hagoort, P., & Cutler, A. (2012). Rapid recognition at 10 months as a predictor of language development. *Developmental Science, 15*(4), 463–473. http://dx.doi.org/10.1111/j.1467-7687.2012.1144.x

Newman, R., Bernstein Ratner, N., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: A retrospective analysis. *Developmental Psychology, 42*(4), 643–655. http://dx.doi.org/10.1037/0012-1649.42.4.643

Singh, L., Reznick, J. S., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: A longitudinal analysis. *Developmental Science, 15*(4), 482–495. http://dx.doi.org/10.1111/j.1467-7687.2012.01141.x

van de Weijer, J. (1998). *Language input for word discovery* (Vol. 9). Nijmegen, The Netherlands: Max Planck Series in Psycholinguistics.