

Children's Use of Disfluencies for Pragmatic Inference in Lexical Development

Celeste Kidd, Katherine S. White, and Richard N. Aslin

{ckidd, kwhite, aslin}@bcs.rochester.edu

Department of Brain & Cognitive Sciences, Meliora Hall

Rochester, NY 14627 USA

Abstract

In early lexical development, children must learn to map spoken words onto their respective referents. Since multiple objects are typically present when any word is used, a child is charged with the difficult task of inferring the speaker's intended referent. Previous research has uncovered various cues children may use in this task, including contextual and social cues. We investigate a previously unexplored cue for inferring speaker intention during word learning: speech disfluencies. Disfluencies (such as "uh" and "um") occur in predictable locations, such as before words that are infrequent and words that have not been previously mentioned. We conducted an eye-tracking study to investigate whether young children can make use of the information contained in disfluencies to infer a speaker's intended referent. Our results demonstrate that young children (ages 2;4 to 2;8) are sensitive to disfluencies. More critically, they show that children appear to *use* disfluencies predictively as a cue to reference and to speaker intention as the disfluency is occurring. We also examined potential sources of learning about disfluencies in CHILDES (MacWhinney, 2000) and found that disfluencies, though rare, occur regularly and with increasing frequency over time in child-directed and child-produced speech. These results reveal that young children attend to speech disfluencies relatively early in lexical development and are able to use the disfluencies to infer speaker intention during online comprehension.

Keywords: Language development; lexical development; word learning; pragmatic inference; speech disfluency.

Introduction

Word learning is fundamental to language acquisition, but the successful mapping of auditory events (spoken words) onto referents (objects in the world) is not always transparent. While some labeling contexts are unambiguous (e.g., holding a cookie and saying "cookie"), most contexts involve multi-word utterances and multiple objects in the child's visual field. Thus, extra-linguistic cues, such as inferring speaker intention, can play a crucial role in word learning. The inability to correctly infer speaker intent, as in autism, can lead to catastrophic failures of communication (Preissler & Carey, 2005).

Previous work has explored various cues available to learners that could aid in determining speaker intention. Social cues, such as parent-child joint visual attention and pointing, provide cues to word learners about the intended referent (Butterworth, 1980; Southgate, 2002). By 18 months, children are even able to make use of information conveyed by the speaker's eye-gaze (Baldwin, 1991). Discourse context may also aid in determining a speaker's intended referent (Frank et al., 2009).

In addition to these externally available cues, young children appear to make certain assumptions that facilitate rapid lexical development. One assumption of particular relevance is that of mutual exclusivity (Markman, 1990). Experimental evidence suggests that young word learners assume a one-to-one mapping between words and referents, starting as young as 15 months of age (Halberda, 2003; Markman, Wasow, & Hansen, 2003).

Here we investigate a previously unexplored cue for inferring speaker intention: speech disfluencies. Disfluencies (e.g., "uh" and "um") could be a potentially powerful cue to reference for several reasons. First, disfluencies occur in highly predictable locations—specifically, before unfamiliar or infrequent words, and before words that have not previously been mentioned in the prior discourse. Thus, rather than being noise, disfluencies provide *information* which is potentially useful for discovering what a speaker intends to refer to. Since disfluencies occur before an object is labeled, they could enable a child to anticipate upcoming referents and therefore enhance the speed of spoken word recognition.

Disfluencies are known to be a reliable property of speech between adults. Fox Tree (1995) estimated that about 6 disfluencies occur per 100 words, excluding pauses (which are not necessarily disfluencies). Shriberg (1996) found that disfluencies occur every 7 to 15 words in conversation between adults, depending upon which corpus (SWBD or AMEX) was used for the analysis. The rate of disfluency varies as a function of several factors: speaker familiarity, utterance length, and speech rate (Shriberg, 1996).

Disfluencies come in a variety of different types, including repetitions, substitutions, insertions, deletions, and speech errors. We focus here on the most common type of disfluency, the *filled pause*—"uh" and "um" in English (Shriberg, 1996). This type of disfluency is particularly common before infrequent and previously unmentioned words (Arnold & Tanenhaus, 2007). Consider the following example of a filled pause from the Sachs corpus in CHILDES (MacWhinney, 2000):

(1) *CHI: *Telephone?*
*MOT: *No, that wasn't the telephone, honey.*
That was the, uh, timer.

The disfluency occurs before the less frequent word and previously unmentioned object, "timer". Such disfluencies are thought to result from a delay in lexical retrieval (Clark & Fox Tree, 2002; Fox Tree & Clark, 1997).

There is evidence that adults use disfluencies online during sentence comprehension. Arnold and colleagues demonstrated that adults can use disfluencies to anticipate that an upcoming referent is likely to be new to the discourse (Arnold et al., 2004; Arnold, Fagnano, & Tanenhaus, 2003) or less frequent (Arnold, Hudson Kam, & Tanenhaus, 2007). In a series of eye-tracking experiments, adults were biased to look at discourse-new or unfamiliar objects when labels were preceded by a disfluency.

In the present study, we explore whether young children can use disfluencies to infer the identity of an upcoming referent.

Experimental Data

Methods

Participants Sixteen parents from the Rochester community volunteered their toddlers. The parents were recruited through mailings, posted flyers, and web ads. The children ranged in age from 2;4 to 2;8 ($M = 2;6$), had no reported hearing deficits, and were from monolingual, English-speaking homes. Participants received either \$10 or a toy as compensation.

Stimuli The stimuli consisted of 16 pairs of items, each containing one familiar item (e.g., *ball*) and one novel item (e.g., *dax*). The 16 familiar items were selected from among the earliest acquired English words in the MacArthur-Bates Communicative Development Inventories (Dale & Fenson, 1996). The novel items were picked to match the familiar items in visual complexity. Each novel item was assigned a novel word. Novel words were matched to the familiar words in syllable length, word onsets, and stress patterns.

Apparatus Eye-tracking was performed using a table-mounted Tobii 1750 eye-tracker with a 17-inch monitor. The stimuli were presented using Psychopy running on a Mac Mini with an Intel Core 2 Duo processor.

Procedure Each child was seated on a parent's lap with the child's eyes approximately 23 inches from the Tobii monitor. The parent wore headphones playing music to mask the auditory stimuli and prevent influence on the child's behavior. The experiment consisted of 16 trials, each preceded by an attention-getter in the center of the screen. For each trial, an object pair was displayed (e.g., a *ball* and a *dax*). Whether the novel item appeared on the left or right was balanced across the entire experiment. Each trial consisted of three phases: two discourse phases and one critical phase. In the first discourse phase, the object pair appeared and 500 ms later a recorded voice introduced the familiar object (e.g., "I see the *ball*!"). After 1,000 additional ms, the object pair disappeared. The second discourse phase began 500 ms later with the reappearance of the same object pair. The discourse again referred to the familiar object (e.g., "Ooooh! What a nice *ball*!"). The objects disappeared 1,000 ms after the recorded voice ended.

During the third critical phase, children were instructed to look at one of the two objects (familiar or novel) with either a fluent or a disfluent command (Table 1). Whether the target object was familiar or novel, and whether the command was fluent or disfluent, was balanced throughout the experiment, such that four trials of each type occurred for each child. The assignment of particular item pairs to each condition was counterbalanced across participants.

Table 1: Trial type examples.

	Familiar target	Novel target
Fluent	<i>I see the ball!</i>	<i>I see the dax!</i>
Disfluent	<i>I see thee, uh, ball!</i>	<i>I see thee, uh, dax!</i>

If young children are able to make use of the information contained in a disfluency to anticipate that an upcoming referent is likely to be novel and/or previously unmentioned, we would expect to see more looks to the novel object during the period of disfluency. Thus, we used looks to the novel object during the 2-second window prior to the onset of the target word as our dependent measure. (In the disfluent trials, this is the period of time when the disfluency is occurring.) The object pair appeared at the beginning of this 2-second window and remained on-screen for 5 seconds.

In the disfluent trials, the earliest sign of the disfluency is at the determiner—"thee" in disfluent trials versus "the" in fluent speech. Thus, the determiner was chosen as the onset of the window of interest. Because our focus was on anticipatory looking to the target, the window of interest ended at the onset of the target word.

Due to the nature of disfluencies, the disfluent command took longer to execute; consequently, the linguistic material in the window of interest varied across fluent and disfluent trials. To compensate for this difference, the command "Look!" was repeated in all trials. Thus, in all trials, children had been instructed to look at the screen before the object pair reappeared and the window of interest began. The first "Look!" instruction was successful in directing children's attention to the screen: on 88.4% of trials, children were looking at the screen immediately prior to the onset of the window of interest.

Results

To ensure that children looked reliably at the appropriate object after it was named, we calculated for each trial type (fluent, disfluent) the proportion of time the child looked at the target item during the 2 seconds after the target was named. On trials in which the target was familiar, the mean proportion of looking to the target during this window was 75.6%. Thus, children reliably mapped familiar words to familiar objects. During trials in which the target was novel, the mean proportion of looking to the target was 75.2%, demonstrating that children mapped novel words to novel objects as reliably as they mapped familiar words to familiar objects. Taken together, these results show that children

consistently arrived at the target object, regardless of the trial type.

Next, we calculated the proportion of looks to the novel object at each time point during the critical phase of the fluent and disfluent trials. Figure 1 shows the resulting timecourse plot for trials in which the target was novel. As predicted by our hypothesis, children looked more towards the novel object during the 2-second window of interest (*before* the onset of the target word) in the disfluent trials than they did in the fluent ones. This suggests that children were able to make use of the information contained in a disfluency to anticipate an upcoming novel referent. Figure 2 shows the timecourse plot for trials in which the target was the *familiar* object. In these trials also, children looked more towards the novel object during the window of interest in the disfluent trials than in fluent ones. This suggests that the disfluency set up the expectation of an upcoming novel referent, though that expectation was erroneous in these *familiar*-object trials. However, both Figures 1 and 2 show that, after the onset of the target word, children correctly identify the target picture.

The timecourse plots suggest that children were sensitive to the presence of the disfluency and were biased to interpret that disfluency as signaling that the upcoming word would refer to the novel/previous unmentioned referent. To test that hypothesis, we compared looks to the novel object across fluent and disfluent trials in the 2-

second window of interest before the onset of the target word. During disfluent trials, children looked at the novel object for 1158 ms. During fluent trials, children looked at the novel object for 893 ms. A Wilcoxon signed-rank test found this difference to be highly significant ($p < 0.008$). This result suggests that children are sensitive to disfluencies and use them predictively to infer that an upcoming referent is likely to be novel and/or previously unmentioned.

However, an alternative explanation for this result is that children simply paid more attention overall to the display (both objects) during disfluencies. To further examine whether disfluencies cause preferential looking to the novel object, we compared the average *proportion* of total looking time to the novel object during the same temporal window of interest. Children looked at the novel object 66% of the time in the disfluent trials, as opposed to 54% of the time in the fluent trials. A Wilcoxon signed-rank test found this difference to be highly significant ($p < 0.005$). Further, the proportion of looking time to the novel object was significantly above chance in the disfluent trials ($p < 0.001$), whereas in the fluent trials, children's looking to the two objects did not differ from chance ($p > 0.37$). These results demonstrate that disfluencies cause a selective increase in attention to the novel objects, suggesting that children use disfluencies online to create expectations about the speaker's intended referent.

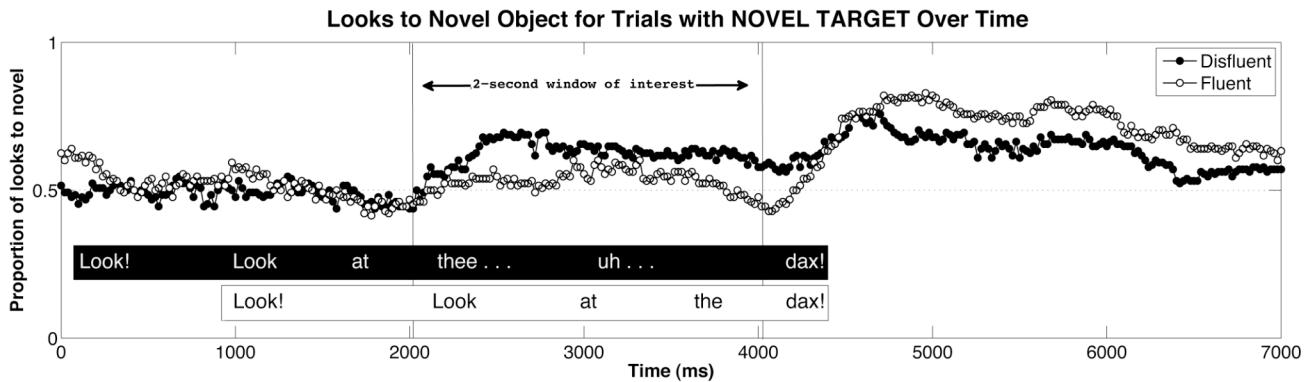


Figure 1: The proportion of looks to the novel object over time for the critical phase of the trials with novel targets.

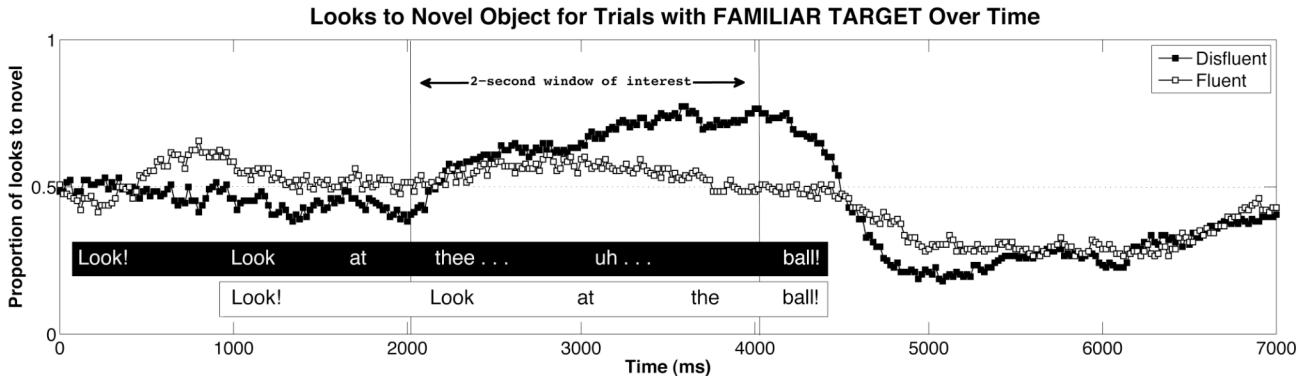


Figure 2: The proportion of looks to the novel object over time for the critical phase of trials with *familiar* targets.

Disfluencies in child-directed and child-produced speech

The results of our eye-tracking study indicate that young children learn that speech disfluencies contain information about upcoming referents, but, at present, the source of this learning is unknown. As discussed above, disfluencies are abundant in speech among adults. Young children's demonstrated sensitivity to disfluencies, then, could indicate that they learn about disfluencies from adult-directed speech. In fact, some experimental evidence demonstrates that young children (1;10) attend to disfluencies in speech between adults (Soderstrom & Morgan, 2007). This hypothesis may seem particularly attractive given that child-directed speech is characterized by a slow speech rate and short utterances, both of which correlate with more fluent speech. This fact may explain why fluency is widely regarded as a hallmark of child-directed speech.

The true frequency with which disfluencies occur in child-directed speech, however, is unclear. Although child-directed speech is undeniably more fluent than adult-directed speech, disfluencies occur in speech to children more than is typically recognized. An informal search of CHILDES (MacWhinney, 2000) revealed several types of disfluencies involving filled pauses in child-directed speech. Consider the following examples:

- (2) *MOT: *Should I take the, uh, bologna ?*
- (3) *MOT: *Well are you gonna marry Batman—uh, Robin? Or is Donna going to marry Robin?*
- (4) *MOT: *No, that's not Fozzie, that's, uh... Fozzie's the bear.*

The above examples from child-directed speech represent many of the types of disfluencies reported to occur in adult-directed speech. Example 2 contains a simple filled pause before an infrequent word, *bologna*. Example 3 contains a filled pause followed by a substitution for an incorrect word, *Batman*. Example 4 contains a filled pause (presumably produced while the mother attempts to retrieve an obscure Muppet's name) followed by an abandonment of the first sentence (sometimes called a "deletion" in the adult literature on disfluencies). If such disfluencies are a reliable property of child-directed speech, they could serve as a sufficient source of learning for children.

An analysis of child-directed speech in CHILDES was conducted to evaluate whether disfluencies are a reliable property of speech to children. Only transcripts containing two participants (target child and adult caretaker) were included to ensure that extracted disfluencies were directed at the target child, and not to an older sibling or another adult. The average probability of a filled pause was calculated for speech to children at each age.

Figure 3 (upper) demonstrates the regularity with which filled pauses occur in speech to young children. At age 2,

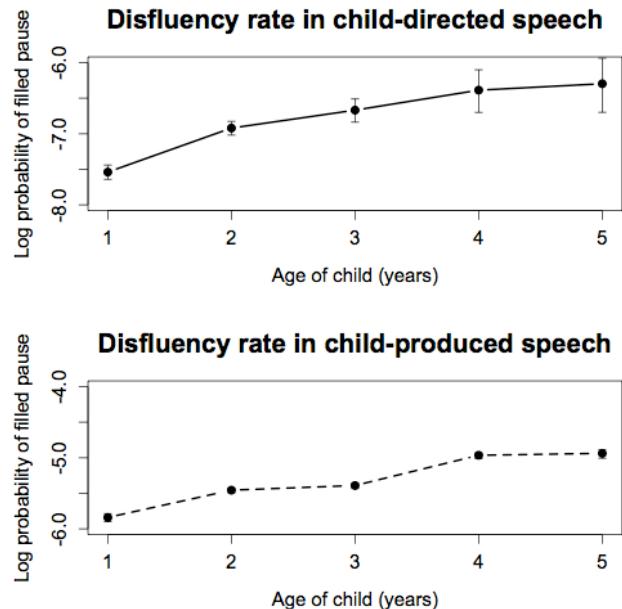


Figure 3: Log probability of a filled pause in child-directed speech (upper) and child-produced speech (lower) in CHILDES, plotted by age with standard errors.

filled pauses occur at an approximate rate of 1 every 1,000 words. For comparison, Shriberg (1996) estimated the rate to be 1 every 50 words in adult-directed speech from the Switchboard corpus. Further, the plot demonstrates that children hear filled pauses more frequently as they get older (Spearman's rank correlation¹, rho = 0.85, p < 0.002), in accord with the fact that caretakers tend to use longer, more complicated utterances with older children.

One important caveat is that only transcribed filled pauses can be detected in the CHILDES transcripts. Since these parent-child interactions were not transcribed specifically for the purpose of analyzing speech disfluencies (with the exception of the Soderstrom corpus), it is likely that transcribed disfluencies represent only a subset of those that occurred. Thus, this analysis may not accurately reflect the absolute rate. It does, however, suggest that disfluencies are a reliable feature of speech to children, and that disfluencies become increasingly more frequent with age.

Thus far, we've established that children could learn the distributional properties of disfluencies from either adult- or child-directed speech. We would like to address one more possibility—namely, that children gain an understanding of disfluencies through their own productions of disfluencies. Consider the following child-produced disfluencies from CHILDES:

- (5) *CHI (2,5): *This, uh, this ...*
- *MOT: *What? Show me .*
- *CHI: *This, uh, this—uh, this.*

¹ In this type of analysis, the value of rho depends on the bin size. To ensure enough data was available for each bin, we used a bin size of 6 months.

- *MOT: *Show me.*
- *CHI: *This. Uh, this!. This! This!*
- *MOT: *What is it? What is he looking at ?*

- (6) *CHI (3;0): *I want some breakfast? Uh, de, uh, uh, de—I, de—I get two breakfasts out .*
- (7) *CHI (3;2): *I wanna get some candy from—from Scotty for Valentine.*

In the above examples, children produced disfluencies in association with presumed lexical retrieval difficulties. Figure 3 (lower) shows that the probability of a filled pause increases in child-produced speech from CHILDES with age² ($\rho = 0.94$, $p < 0.001$). At age 2, the rate is approximately 1 every 230 words. Thus, a child's understanding of disfluencies may also change as a result of detecting regularities in her own productions.

Discussion

Many contemporary theories model word learning as a process of learning the arbitrary association between sounds and meanings (Frank, Goodman, & Tenenbaum, 2008; Siskind, 1996; Yu & Ballard, 2007). The results of our experiment demonstrate that young children's ability to match sounds with meanings is considerably more general: they are able to match disfluencies with a speaker's intended referent, a property of communication that is not directly observed. These results raise several important issues.

First, it is unclear whether novelty or discourse status is driving these effects. Adults' interpretation of disfluencies is affected by both of these factors. In our study, the novel objects were both previously unmentioned and novel. Work in progress attempts to uncover which of these—or both—drives the effect.

Second, what children understand about disfluencies is an open question. Clark and colleagues have suggested that speech disfluencies signal to the listener that a speaker is having difficulties producing speech (Clark & Fox Tree, 2002). Furthermore, there is some evidence that adults respond to disfluencies in part because they understand that the disfluency is driven by the speaker's processing difficulties. For example, Arnold, Hudson Kam, and Tanenhaus (2007) demonstrated that adult listeners do not use disfluencies predictively to infer an upcoming referent if they are told the speaker has a type of brain damage that causes disfluent speech.

It is possible that children, too, engage in this type of causal reasoning. Children may be aware that disfluencies are the result of processing (specifically, lexical access) difficulties, and therefore look for a referent that is likely to have caused difficulties. While this reasoning almost certainly does not happen consciously, children may nonetheless have learned that disfluencies occur because of

speaker difficulty, and that speaker difficulty often arises with novel referents. If so, we might expect children—like adults (Arnold et al., 2007)—to alter their interpretations when disfluencies can be attributed to an external cause.

Of course, children could potentially show the patterns demonstrated here without any explicit understanding of the linguistic processing mechanisms involved. Disfluencies might simply be associatively linked through experience to novel referents. That is, disfluencies could be treated just like words that mean “look at the novel referent”. This theory does not assume intermediate stages of processing or conceptual reasoning: the association between a disfluency and a meaning is direct and quick. Such an account would predict that children could not alter their interpretation of disfluencies based on whether they were perceived as internally or externally driven.

Both accounts are plausible, given what is known about infants' and young children's capabilities. Infants and children are known to be capable statistical learners (e.g., Fiser & Aslin, 2001; Saffran, Aslin, & Newport, 1996), which could enable them to detect correlations between disfluencies and referent novelty in the environment. Young children are also able to engage in pragmatic inference (e.g., Behne, Carpenter, & Tomasello, 2005) and even very young children are able to infer the intentions and difficulties of others (Warneken & Tomasello, 2006). Therefore, it is possible that children have access to this type of reasoning during online sentence processing.

If children learn about disfluencies from their environment, does their environment provide them with enough data to detect these relationships? Our corpus analyses demonstrate that disfluencies are a reliable feature of speech to young children. Since disfluencies are prosodically salient, even infrequent occurrences could provide young children with enough evidence to learn about the information they convey. Prosodic information is available to infants extremely early—even before lexical boundary information. In fact, acoustic analyses by Soderstrom and colleagues (in press) find that disfluencies in speech to infants and children are exceptionally prosodically salient—longer and higher in pitch than those that occur in speech to adults. Thus, a child could learn the distributional properties of disfluencies by accumulating evidence from infrequent examples over time. Of course, children could also learn about disfluencies by attending to speech among adults.

While statistical learning is one potential mechanism by which children could learn to use disfluencies in processing, it is also possible that they learn to use them from their own productions. The results of the CHILDES search above showed that even young children produce disfluencies. Thus it is possible that children gain the understanding that disfluencies signal processing difficulty by producing disfluencies themselves.

Together, the results of this study indicate that young children (1) have learned that disfluencies contain information, (2) attend to disfluencies in speech, and (3) can

² As in the earlier analysis on child-directed speech, ρ was calculated using a bin size of 6 months.

make use of the information contained in disfluencies online during comprehension in order to infer speaker intention.

Acknowledgments

The first author was supported by a National Science Foundation Graduate Research Fellowship. The research was supported by a National Institutes of Health grant (HD-37082). We thank Johnny Wen and Steven T. Piantadosi for assistance with analyses; Suzanne Horwitz, C. Michael Lindsey, Kristen Graziano, Fahria Omar, Emily Richards, Lee Kehoe, Lindsay Woods, and Rosemary Ziemnik for recruiting and scheduling subjects; and Jennifer Arnold, Michael Tanenhaus, Elissa L. Newport, Mohinish Shukla, Casey Lew-Williams, Anne Pier Salverda, Michael C. Frank, Melanie Soderstrom, Eve V. Clark, Jesse Snedeker, Herbert H. Clark, Michael S. DeFreitas, and Meredith Brown for advice, comments, and suggestions.

References

Arnold, J. E & Tanenhaus, M. K. (2007). Disfluency effects in comprehension: how new information can become accessible. In E. Gibson & N. Pearlmuter (Eds.), *Referential processing in adults and children*. Cambridge, MA: MIT Press.

Arnold, J. E., Tanenhaus, M. K., Altmann, R. J., & Fagnano, M. (2004). The Old and Thee, uh, New. *Psychological Science*, 9, 578-582.

Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies Signal Theee, Um, New Information. *Journal of Psycholinguistic Research*, 32(1), 25-36.

Arnold, J. E., Hudson Kam, C. L., & Tanenhaus, M. K. (2007). If You Say *Thee uh* You Are Describing Something Hard: The On-Line Attribution of Disfluency During Reference Comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 914-930.

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, 63, 975-890.

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science*, 8, 492-499.

Butterworth, G. & Cochran, E. (1980). Towards a Mechanism of Joint Visual Attention in Human Infancy. *International Journal of Behavioral Development*, 3(3), 253-272.

Clark, H. H. & Fox Tree, J. E. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition*, 84(1), 73-111.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, 28, 125-127.

Fiser, J. & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, 12, 499-504.

Frank, M. C., Goodman, N. D., & Tenenbaum, J. (2008). A Bayesian framework for cross-situational word learning. *Advances in Neural Information Processing Systems*, 20.

Frank, M. C., Goodman, N. D., Tenenbaum, J. B., & Fernald, A. (2009). Continuity of Discourse Provides Information for Word Learning. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 31st Annual Cognitive Science Society*. Amsterdam, the Netherlands: Cognitive Science Society.

Fox Tree, J. J. (1995). The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language*, 34, 709-738.

Fox Tree, J. J., & Clark, H. H. (1997). Producing "the" as "thee" to signal problems in speaking. *Cognition*, 62(2), 151-167.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, 87(1), B23-B34.

MacWhinney, B. (2000). The CHILDES project: Tools for analyzing talk. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, 14(1), 57-77.

Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology*, 47(3), 241-275.

Preissler, M. A., and Carey, S. (2005). The role of inferences about referential intent in word learning: Evidence from autism. *Cognition*, 97(1), B13-B23.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month old infants. *Science*, 274, 1926-1928.

Shriberg, E. E. (1996). Disfluencies in SWITCHBOARD. *Proceedings of the International Conference on Spoken Language Processing, Addendum* (pp. 233-251). Philadelphia, PA.

Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, 61(1), 39-91.

Soderstrom, M., Blossom, M., Foyel, I., & Morgan, J. L. (in press). Acoustical cues and grammatical units in speech to two preverbal infants. *Journal of Child Language*.

Soderstrom, M. & Morgan, J. L. (2007). Twenty-two-month-olds discriminate fluent from disfluent adult-directed speech. *Developmental Science*, 10, 641-653.

Southgate, V., van Maanen, C., & Csibra, G. (2007). Infant Pointing: Communication to Cooperate or Communication to Learn? *Child Development*, 78(3), 735-740.

Warneken, F. & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, 31, 1301-1303.

Yu, C. & Ballard, D. A unified model of word learning: Integrating statistical and social cues. *Neurocomputing*, 70(13-15), 2149-2165.