Function Words in Early Speech Perception

Rushen Shi[†], Janet Werker[‡] and Anne Cutler[*]
† Université du Québec à Montréal, Canada
‡ University of British Columbia, Canada
* Max Planck Institute for Psycholinguistics, The Netherlands
E-mail: shi.rushen@uqam.ca, jwerker@psych.ubc.ca, anne.cutler@mpi.nl

ABSTRACT

Three experiments examined whether infants recognise functors in phrases, and whether their representations of functors are phonetically well specified. Eight- and 13-month-old English infants heard monosyllabic lexical words preceded by real functors (e.g., *the, his*) versus nonsense functors (e.g., *kuh*); the latter were minimally modified segmentally (but not prosodically) from real functors. Lexical words were constant across conditions; thus recognition of functors would appear as longer listening time to sequences with real functors. Eight-month-olds' listening times to sequences with real versus nonsense functors did not significantly differ, suggesting that they did not recognise real functors, or functor representations lacked phonetic specification. However, 13-month-olds listened significantly longer to sequences with real functors. Thus, somewhere between 8 and 13 months of age infants learn familiar functors and represent them with segmental detail. We propose that accumulated frequency of functors in input in general passes a critical threshold during this time.

1. INTRODUCTION

Recognising words in continuous speech appears effortless and automatic. Yet the spoken forms of words are variable since sentential and phonetic context, speaker differences, etc., can all affect realisations of a word. Infants face this variability problem when acquiring a language. In this study we examine the development of function word (functor) recognition in English-learning preverbal infants. Functors include structurally important words such as determiners, auxiliaries, and complementisers. In contrast to lexical words (nouns, verbs, adjectives and adverbs), which are more complex acoustically and phonologically in input speech to children, functors are minimised in their spoken form across languages [1, 2]. This raises the question of when infants begin to recognise these words and how they represent these word forms.

Recent research has shown that during the second half of the first year infants begin to recognise some word forms [e.g., 3, 4, 5]. The fact that these word forms are typically lexical words is consistent with the evidence that even newborns are sensitive to the spoken form distinction between lexical and function words [6] and that by 6 months of age infants additionally show a listening preference for the acoustically and phonologically more complex lexical words [7, 8]. These findings do not necessarily mean that preverbal infants fail to recognise functors. After all, given that functors occur far more frequently in speech than lexical words, it is reasonable that infants may recognise them quite early. On the other hand, unlike lexical words, functors are more likely to be reduced and cliticised in spontaneous utterances [9, 10]. Furthermore, they hardly ever occur in citation forms. This suggests that the nature of the recognition and of the phonetic representation for function words in infancy may differ from that for content words.

One previous study reported that 7- to 15-month-old German-learning infants can recognise functors [11]. Shady [12], also using the headturn preference procedure, showed that 10.5-month-old English-learning infants can detect functors. The stimuli that Shady used were synthesised sentences containing either unmodified or modified functors. To date, no other research group has conducted work on infants' perception of functors. In the present study, naturally produced speech samples and a different procedure (Central Visual Fixation Procedure) were used to examine monolingual English-learning infants' recognition and representation of these words.

2. EXPERIMENT 1

In this experiment we asked if 8-month-old infants are able to recognize functors in a simple noun phrase context (such as *the* in *the chair*). Infants were presented with such phrases versus phrases with segmentally altered nonsense functors[i]. Functor recognition should be indicated by longer listening time for phrases with real functors than for phrases with nonsense functors. This method of assessing functor recognition is appropriate due to the clitic nature of functors in natural speech [10].

Stimuli: Stimuli were 5 real English functors, *the, his, her, their,* and *its*, and 5 nonsense functors, *kuh, ris, ler, lier*, and *ots*. The nonsense functors were different from the real functors segmentally but not prosodically. Thus, *kuh, ris, ler* and *ots* differ from *the, his, her* and *its* respectively only in the initial segment; *lier* differs from *their* in the initial consonant and the following vowel ([I] for *lier*) but not in the final retroflexed schwa. In addition, 2 lexical non-words, *tink* [tInk] and *doomp* [dump] were created. There were 2

 ISBN 1-876346-48-5

conditions in this experiment. One condition included noun phrases consisting of a real functor followed by *tink* versus phrases consisting of a nonsense functor followed by *doomp*, i.e., "RealFunctor + *tink*" (e.g., *the tink, their tink, her tink*, etc.) versus "NonsenseFunctor + *doomp*" (e.g., *ris doomp, lier doomp*, etc.). The real functors each had comparable chances to co-occurr with *tink*, with randomized phrase orders. The nonsense functors were assigned to *doomp* in the same fashion. The second condition included phrases consisting of a real functor followed by *doomp* versus phrases consisting of a nonsense functor followed by *tink*, i.e., "RealFunctor + *doomp*" versus "NonsenseFunctor + *tink*". The real functors each had comparable chances to co-occur with *doomp*, and the nonsense functors each had comparable chances to co-occur with *tink*. The phrases were produced in infant-directed register by a native English speaker, a mother who had a 6-month-old infant at the time of recording. She was blind to the purpose of the study.

Participants: Sixteen 8-month-old monolingual English-learning infants completed the task. Half of these infants were tested in the first condition (Group 1) and the other half in the second condition (Group 2).

Procedure and design: the Central Visual Fixation Procedure [13] was used to assess infants' recognition of real functors. The infant was seated on the parent's lap in front of a TV screen and loudspeaker. During trials, the auditory stimuli were presented together with a visual display of a black-and-white checkerboard. The parent wore headphones delivering masking music. Each 16-second trial was initiated upon the infant's fixation. His/her total looking time to the screen during each trial was recorded online. The experimenter in an adjacent room, who was blind to the trial orders, observed the infant's eye movement through a closed-circuit TV and pressed a computer key whenever an eye fixation occurred. Each test session was recorded to allow for post-test coding, necessary for the verification of the online coding. Six trials were presented. For the first condition (Group 1), 3 "RealFunctor + *tink*" trials alternated with 3 "NonsenseFunctor + *doomp*" trials, with order of trials counterbalanced. Likewise for the second condition (Group 2), 3 "RealFunctor + *doomp*" trials alternated with 3 "NonsenseFunctor + *tink*" trials. Each trial contained multiple phrases of the same kind, e.g., a trial of "RealFunctor + *tink*" could contain phrases such as *the tink, her tink, their tink, his tink, etc.*

Results: To assess infants' recognition of function words in phrases, we compared each infant's looking time while listening to "RealFunctor + LexicalWord" phrases with his or her looking time while listening to "NonsenseFunctor + LexicalWord" phrases. For each infant, we calculated the total looking time across "RealFunctor + LexicalWord" and "NonsenseFunctor + LexicalWord" trials separately. A mixed 2x2 ANOVA was performed with Group (Group 1 versus Group 2) as the between-subjects factor and phrase type (real functor phrases versus nonsense functor phrases) as the within-subjects factor. There were no significant main effects of Group or of Phrase type, and no interaction (M real functor phrases =37.88 secs, SE=1.49; M nonsense functor phrases = 38.48 secs, SE = 1.23). We interpreted these results as a lack of evidence for real functor recognition in phrases in 8-month-old infants.

One concern of this experiment is the appropriateness of the lexical non-word *doomp* [dump], since tense vowels do not occur in the environment of /_mp/ in English. We therefore conducted a subsequent experiment replacing *doomp* with a more English-like word form *breek* [brik].

3. EXPERIMENT 2

This experiment replicates Experiment 1, with the lexical non-word *doomp* replaced by another non-word *breek* [brik], a word form more in accord with English phonology. Our goal is to confirm if 8-month-olds recognize function words in a simple noun phrase context.

Stimuli: The stimuli for Experiment 2 are identical to those of Experiment 1, except for the replacement of *doomp* with *breek*. Again, there were two conditions, one involving "RealFunctor + *tink*" versus "NonsenseFunctor + *breek*" phrase types, and another involving "RealFunctor + *breek*" versus "NonsenseFunctor + *tink*" phrase types.

Participants: Sixteen 8-month-old monolingual English-learning infants completed this experiment. Eight took part in each condition.

Procedure and design: Identical to Experiment 1, except that *doomp* was replaced by *breek*.

Results: As in Experiment 1, infants' recognition of function words was assessed by comparing each infant's looking time while listening to "RealFunction + LexicalWord" phrases with his or her looking time while listening to "NonsenseFunctor + LexicalWord" phrases. As in Experiment 1, a mixed 2x2 ANOVA was performed with Group (Group 1 versus Group 2) as the between-subjects factor and Phrase type (real functor phrases versus nonsense functor phrases) as the within-subjects factor. Again, there were no significant main effects and no significant interaction (M real functor phrases =37.80 secs, SE=1.15; M nonsense functor phrases = 37.81sec, SE = 1.17. See Figure 1). Therefore, even when the word forms of the stimuli accorded with English phonology, there was still no evidence of real functor recognition in phrases in 8-month-old infants. Thus we next asked whether infants could recognize these real functors at a later stage.

4. EXPERIMENT 3

The purpose of this experiment was to determine whether 13-month-old infants recognize function words in simple noun phrases. We chose the age of 13 months because by this time infants are significantly more advanced in language acquisition. They may recognize up to 25 spoken words, and may be beginning to produce words. Moreover, it has been suggested in previous work that as early as 11

months of age infants show a differential ERP response to real versus nonsense functors in ongoing speech [14]. As in Experiment 1 and 2, infants were presented with phrases containing a real functor followed by a lexical word versus phrases containing a nonsense functor followed by another lexical word. If they recognise the functors, we would expect them to look longer while listening to phrases containing real functors versus nonsense functors.

Stimuli: Identical to those of Experiment 2.

Participants: Sixteen 13-month-old monolingual English-learning infants. As in Experiment 2, eight took part in each condition.

Procedure and design: Identical to Experiment 2.

Results: As in Experiment 2, infants' recognition of function words was assessed by comparing each infant's looking time while listening to "RealFunction + LexicalWord" phrases with his or her looking time while listening to "NonsenseFunctor + LexicalWord" phrases. A mixed 2x2 ANOVA was performed with Group (Group 1 versus Group 2) as the between-subjects factor and Phrase type (real functor phrases versus nonsense functor phrases) as the within-subjects factor. The results showed that there was a significant main effect of Phrase type, $F(1, 14) = 10.691$, $p = .006$. ($M=36.46$ secs, $SE=1.29$ secs for real functor phrases; $M=32.84$ secs, $SE=1.51$ secs for nonsense functor phrases), but no significant main effect for Group and no interaction (again, see Figure 1). Therefore, our results indicate that 13-month-olds recognised real functors. In addition, as the nonsense functors in our study were minimally modified from real functors, the results suggest that infants not only recognized real functors, but also represented these words with detailed segmental specifications.

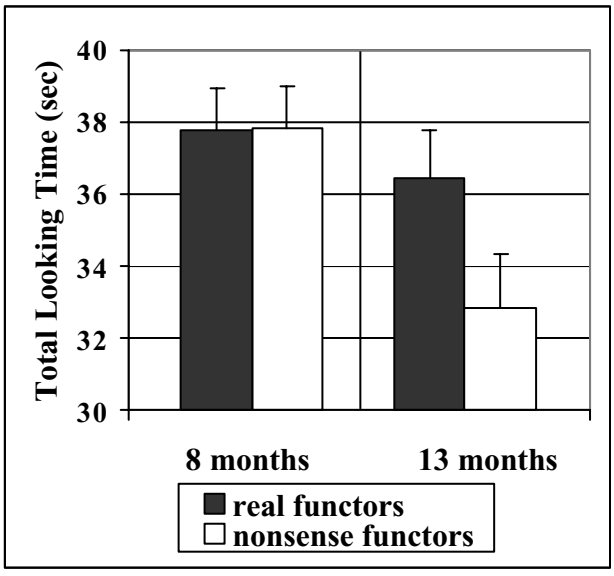

Figure 1: Mean total looking time for infants of 8 months (Experiment 2) and infants of 13 months (Experiment 3) to phrases with real versus nonsense functors.

5. GENERAL DISCUSSION

In three experiments we showed that 8-month-old English-learning infants did not treat function words in simple noun phrases differently from nonsense functors, whereas 13-month-old infants did. These results suggest that the recognition of function words begins to occur somewhere between 8 and 13 months of age. Furthermore, given that the manipulations of our stimuli involved minimal segmental modifications of real functors into nonsense functors, we demonstrated that 13-month-olds not only recognized functors, but also represented them with detailed segmental specifications. We are conducting more research to pinpoint the age at which the recognition for these function words emerges and to determine how their phonetic representations evolve.

In our earlier work infants preferred to listen to the acoustically and phonologically more complex lexical over function words when these two types of words were presented as separate word lists [7, 8]. In the present study we showed that when the co-occurring lexical words were held constant, 13-month-old infants showed a preference for phrases containing familiar real functors over those containing unfamiliar nonsense functors. This pattern of response indicates recognition of real functors.

One factor pertinent for our results concerns the frequency of the function words. Among the 5 functors used in our study, *the* is highly frequent in parental input to preverbal infants, *his* and *her* are far less frequent, with *their* and *its* very infrequent (based on the Brent Corpus, [15]). It is possible that the lack of recognition by 8-month-olds in our experiments was due to the dominant infrequent functors in our stimuli. Infants treated the unfamiliar, infrequent real functors the same as the nonsense functors. The 13-month-olds, on the other hand, recognized the real functors as a whole even though some of them are infrequent. We propose that this is because even the infrequent function words in our stimuli have reached a critical threshold in input when infants are between 8 and 13 months of age. That is, the accumulated frequency of function words in general over a significant period of time has topped the threshold at some point, allowing recognition of functors as a whole to happen. We suggest that this factor is crucial for function word recognition, but not for lexical word recognition.

There are several reasons underlying the above arguments. The recognition of function words in some sense should be more difficult than that of lexical words. Semantically, function words have less transparent word-to-world mapping, as compared with many lexical words. Acoustically and phonologically, function words are universally minimized in spoken form and thus may be less salient than lexical word forms. Moreover, function words rarely occur in citation form, whereas lexical words sometimes do. However, function words as a closed class are far more frequent in occurrence than are lexical words in input speech. Thus, it is plausible that the recognition of function words relies far more heavily on the frequency

ISBN 1-876346-48-5

advantage of these words, as opposed to the case of lexical words. Further detailed studies of large corpora of speech input to preverbal infants as well as experiments with infants testing different frequency factors are necessary to resolve this interesting and complex issue.

Even though the semantic and syntactic features of function words are likely to be acquired much later, the early recognition of the word forms of functors may play an important role in other early language processes such as syntactic analyses (e.g., initial derivation of phrases and basic syntactic classes), lexical word recognition, and the construction of a lexicon.

6. CONCLUSIONS

This study has revealed that between 8 months and 13 months of age English-learning infants begin to recognize function words in continuous phrases, and furthermore, that their representations of these words are phonetically detailed. We suggest that the recognition of functors reflects the attainment of threshold frequency of these words in the input, a factor more crucial for the recognition of function words than for that of lexical words.

7. ACKNOWLEDGEMENTS

This research was supported by a Peter Wall Institute for Advanced Studies Major Thematic Grant on Acoustic Ecology, and by Max Planck Society research stipends, to each of the first two authors, and by a NSERC Discovery Grant to the second author. We thank Daniel Swingley for help with obtaining word frequencies.

REFERENCES

[1] R. Shi, *Perceptual correlates of content words and function words in early language input* (Ph.D. dissertation, Brown University, 1995). *DAI,* 56, 3108A, 1996.

[2] R. Shi, J. Morgan, & P. Allopenna, "Phonological and acoustic bases for earliest grammatical category assignment: A cross-linguistic perspective", *Journal of Child Language,* 25*,* 169-201, 1998.

[3] P. Jusczyk and R. Aslin, "Infants' detection of the sound patterns of words in fluent speech", *Cognitive Psychology,* 29, 1-23, 1995.

[4] P. Jusczyk, D. Houston, & M. Newsome, "The beginnings of word segmentation in English-learning infants", *Cognitive Psychology,* 39, 159-207, 1999.

[5] D. Mandel, P. Jusczyk, & D. Pisoni, "Infants' recognition of the sound patterns of their own names", *Psychological Science*, 6, 315-318, 1995.

[6] R. Shi, J. Werker, & J. Morgan, "Newborn infants' sensitivity to perceptual cues to lexical and grammatical words", *Cognition,* 72, B11-B21, 1999.

[7] R. Shi, and J. Werker, "Six-month-old infants' preference for lexical words", *Psychological Science,* 12, 70-75, 2001.

[8] R. Shi and J. Werker, "The basis of preference for lexical words in six-month-old infants", *Developmental Science*, in press.

[9] A. Cutler, "Phonological cues to open- and closed-class words in the processing of spoken sentences", *Journal of Psycholinguistic Research*, 22, 109-131, 1993.

[10] E. Selkirk, "The prosodic structure of function words", in *Signal to Syntax*, J. Morgan and K. Demuth, Eds, pp. 187-213. Hillsdale, NJ: Lawrence Erlbaum Associates, 1996.

[11] B. Hoehle and J. Weissenborn, "Sensitivity to closed-class elements in preverbal children", in *Proceedings of the 22nd Annual Bonston University Conference on Language Development,* Vol. 1, A. Greenhill, M. Hughes, H. Littlefield, & H. Walsh, Eds, pp. 348-359. Somerville, MA, 1998.

[12] M. Shady, *Infants' sensitivity to function morphemes*. Ph.D. dissertation, SUNY at Buffalo, 1996.

[13] R. Cooper and R. Aslin, "Preference for infant-directed speech in the first month after birth", *Child Development*, 61, 1584-1595, 1990.

[14] V. Shafer, J. Shucard, D. Shucard, & L. Gerken. "An electrophysiological study of infants' sensitivity to the sound patterns of English speech", *Journal of Speech and Hearing Research*, 41, 874-886, 1998.

[15] M. Brent and J. Siskind. "The role of exposure to isolated words in early vocabulary development", *Cognition*, 81, B33-44, 2001.

[i] Following this phase of the experiment, there was another phase presenting different stimuli. Because the second phase was designed to test a different hypothesis, this article will not report that phase.