Developmental Science

Developmental Science 11:3 (2008), pp 407–413 DOI: 10.1111/j.1467-7687.2008.00685.x

PAPER

The effect of functional morphemes on word segmentation in preverbal infants

Rushen Shi and Mélanie Lepage

Département de psychologie, Université du Québec à Montréal, Canada

Abstract

This study examines the role of functional morphemes in the earliest stage of lexical development. Recent research showed that prelinguistic infants can perceive functional morphemes. We inquire whether infants use frequent functors to segment potential word forms. French-learning 8-month-olds were familiarized to two utterance types: a novel noun following a functor, and another novel noun following a prosodically matched nonsense functor. After familiarization, infants' segmentation of the two nouns was assessed in a test phase presenting the nouns in isolation. Infants in Experiment 1 showed evidence of using both frequent functors des *and* mes *(as opposed to the nonsense functor* kes*) to segment the nouns, suggesting also that they had specific representations of the functors. The infrequent functor* vos *in Experiment 2 did not facilitate segmentation. Frequency is thus a crucial factor. Our findings demonstrate that frequent functors can bootstrap infants into early lexical learning. Furthermore, the effect of functors for initial word segmentation is likely universal.*

Introduction

Among the earliest tasks of language acquisition is the segmentation of word forms from continuous speech. Input speech to infants contains primarily multi-word utterances (van de Weijer, 1998; Woodward & Aslin, 1990). Word segmentation is therefore essential for building a lexicon. This task is not easy since there is no obvious cue marking boundaries between words in continuous speech. Furthermore, the acoustic realizations of the same word vary due to factors such as phonetic contexts, speaker variations, and speech rate. Infants must learn to extract potential word forms from larger utterances and categorize variable productions of the same word type.

Since the first demonstration by Jusczyk and Aslin (1995) that English-learning infants can segment words by 8 months of age, numerous studies examined many aspects of word segmentation in infants, including the use of prosodic cues (e.g. Jusczyk, Houston & Newsome, 1999), statistical cues (Saffran, Aslin & Newport, 1996), phonotactic cues (Mattys & Jusczyk, 2001), coarticulation cues (Curtin, Mintz & Byrd, 2001) and the weighting of different cues (Johnson & Jusczyk, 2001; Mattys, Jusczyk, Luce & Morgan, 1999). Moreover, word forms segmented based on statistical learning have been shown to serve as candidates for successful mapping of word meanings (Graf Estes, Evans, Alibali & Saffran, 2007).

Most existing studies on word segmentation were done with English-learning infants using English or artificial languages. Only a few studies examined word segmentation in other languages, showing that infants learning Dutch (Houston, Jusczyk, Kuijpers, Coolen & Cutler, 2000; Kooijman, Hagoort & Cutler, 2004) and Quebec-French (Polka & Sundara, 2003) begin to segment nouns by 8 to 9 months of age. In their goal to understand infants' segmentation strategies, Nazzi and colleagues (Nazzi, Iakimova, Bertoncini, Frédonie & Alcantara, 2006) showed that Parisian infants' word segmentation is syllable based, in contrast to the stress-based strategy used by English-learning infants when they begin to learn to segment words (Jusczyk *et al.*, 1999). Infants thus appear to use language-specific speech cues (e.g. prosody) during early word segmentation. It is likely, however, that language-universal mechanisms may be available at the earliest stage of word segmentation. Cross-language studies provide useful data for better understanding this question.

Function words[1] potentially serve as a language universal cue to word segmentation. This is reasonable, based on several considerations. First, function words and content words in input speech exhibit universal differences in spoken forms across typologically distinct languages (Shi, 1996; Shi, Morgan & Allopenna, 1998). Function words are universally reduced. Even newborn infants can categorically discriminate words of the two broad classes based on language-general differences in

[1] Languages differ with respect to the degree of inflectional complexity. The term 'function words' is used here to include both free and bound functional morphemes. Infants at the early stage of language development may not distinguish between syllabic bound functional morphemes and free functional morphemes.

Address for correspondence: Rushen Shi, Département de psychologie, Université du Québec à Montréal, C.P. 8888, Succursale Centre-Ville, Montréal, Québec, H3C 3P8, Canada; e-mail: shi.rushen@uqam.ca

spoken forms (Shi, Werker & Morgan, 1999). Moreover, in any human language functors are a small set of word types, but each generally occurs highly frequently (e.g. Kucera & Francis, 1967). This contrasts with content words, which are vast in the number of types but each occurs far less frequently. Function words also tend to occur at the edge of phonological phrases, making them potentially more salient to perceive. Perception of function words has been demonstrated in experiments with preverbal infants acquiring English (Shady, 1996; Shafer, Shucard, Shucard & Gerken, 1998; Shi, Werker & Cutler, 2006b), German (Hoehle & Weissenborn, 2003), and French (Hallé, Durand & de Boysson-Bardies, submitted; Shi, Marquis & Gauthier, 2006c). Functors play a role in language comprehension tasks in older infants from about 18 months of age (Bernal, Lidz, Millotte & Christophe, 2007; Fernald & Hurtado, 2006; Gerken, Landau & Remez, 1990; Gerken & McIntosh, 1993; Johnson, 2005; van Heugten & Shi, in press; Kedar, Casasola & Lust, 2006; Zangl & Fernald, 2007) even though these items are typically missing in production at this age. In a one-phase preferential listening task Hallé *et al.* (in press) showed that even 11-month-olds use function words to recognize nouns that they know, and a nonsense syllable that is matched prosodically to a real functor blocks infants' recognition of a familiar noun. These studies show that functors assist lexical processing as soon as infants have a small receptive lexicon.

Can function words bootstrap prelinguistic infants to new word forms? A two-phase task can be used to address this question. A familiarization phase presents a novel word preceded by a functor and another novel word preceded by a nonsense syllable. A subsequent test phase presents both novel words in isolation. Since infants receive the same amount of exposure to the two familiarization stimulus types, a preference during the test phase for the novel word stripped out of the functor context would suggest that infants can use the functor to segment the noun.

In this study we therefore inquire if infants use functors to extract novel vocabulary items and whether frequency is the language universal factor for this effect. Höhle and Weissenborn (2000) showed that preverbal German-learning infants segmented unknown nouns preceded by a determiner, e.g. *der Kahn*. Furthermore, a recent study (Shi, Cutler, Werker & Cruickshank, 2006a) specifically manipulated functor frequencies, revealing that English-learning preverbal infants use frequent functors (such as '*the*'), but not infrequent ones, to segment adjacent nouns. In the present study we extend the previous work to French, using the two-phase task described above. We expect that only frequent functors should assist infants' segmentation of adjacent nouns (e.g. in '*des N*'). It is interesting to examine French for several reasons. In French, a syllable-timing language, functors are less reduced acoustically than in English. Thus, they may be more salient. Moreover, noun phrases (NPs) consistently require a functor before the noun, more so than in English, e.g. *des chiens* (dogs). We inquire if these properties may allow French-learning infants to use frequent functors to segment nouns earlier than English-learning infants. Previous work showed that French-learning infants begin to recognize frequent function words as early as 6 months of age (Shi *et al.*, 2006c). We decided to test infants aged 8 months, an age prior to vocabulary learning.

Experiment 1

Method

Participants

Thirty-two 8-month-old monolingual Quebec-French-learning infants (16 males, 16 females) completed this experiment. Another 12 infants were tested but were excluded from data analyses due to fussiness (eight), parental interference (three) and experimenter errors (one).

Stimuli

Auditory stimuli included a high-frequency French functor *des* (/de/, indefinite plural article), a low-frequency function word *mes* (/me/, '*my*', plural form) and nouns that are rare in infant-directed speech, '*preuve*' ('*proof*') and '*sangle*' ('*girth, saddle*'). Based on Beauchemin, Martel and Théoret (1992), a Quebec-French spoken database (1 million tokens), the frequency of the functors are: *des* 8760, *mes* 1331, *preuve(s)* 50, and *sangle(s)* 0. Analyses of parental speech to the two youngest French-learning infants (a 14-month-old infant from the Champaud transcripts, and a 21-month-old from the Bassano transcripts, combined total tokens 2901) in the CHILDES database (Bassano & Maillochon, 1994; McWhinney, 2000) yielded a similar pattern: *des* 20, *mes* 10, *preuve(s)* 0, and *sangle(s)* 0. The average type–token ratio of all words was 1:8.83 for the Beauchemin database and 1:5.34 for the infant-directed transcripts.

The functors and nouns formed simple NPs – *des preuves*, *des sangles*, *mes preuves*, *mes sangles*. A nonsense functor *kes* (/ke/) that differed from the real functors only in the onset consonant was also used with the nouns to form the following sequences: *kes preuves*, *kes sangles*.

A monolingual Quebec-French female speaker recorded the NPs and the *kes* sequences, and then the nouns in isolation, using the infant-directed speech style. The recording was conducted in an IAC booth, using a DAT recorder, at 44.1 Htz sampling frequency and 16 bits bit rate. The speech was then transferred digital-to-digital to the computer. The final stimuli set consisted of four tokens of each of the six utterances (*des preuves*, *des sangles*, *mes preuves*, *mes angles*, *kes preuves*, *kes sangles*), and four tokens of isolated productions of *preuves* and four tokens of *sangles*.

Design

Infants were randomly assigned to the high-frequency and low-frequency conditions. The 16 infants in the high-frequency condition were familiarized with five trials of functor sequences (i.e. *des+N1* tokens) alternating with five trials of nonsense functor sequences (i.e. *kes+N2*), where N1 and N2 were *preuves* or *sangles*, counter-balanced across infants. Thus, half of these infants heard *des preuves* trials and *kes sangles* trials, and the other half heard *des sangles* trials and *kes preuves* trials. The 16 low-frequency infants were familiarized with five functor trials (i.e. *mes+N1*) alternating with five nonsense-functor trials (i.e. *kes+N2*). Half of these infants heard *mes preuves* trials and *kes sangles* trials, and the other half heard *mes sangles* trials and *kes preuves* trials. Each familiarization trial was 9 seconds long, containing the tokens of a sequence (e.g. *des preuves*) in a random order. One of the four tokens for each sequence was repeated once, yielding five occurrences of a sequence per trial. Across all familiarization trials of each condition, 25 occurrences of the *functor+N* sequences and 25 *kes+N* sequences were presented.

After familiarization, all infants were tested with the isolated nouns, *preuves*, *sangles*, presented respectively in two Test trials. Each Test trial was 18.5 seconds long, containing repetitions of the four tokens of a noun in a randomized order. Half of the tokens of each noun were repeated three times, and the other half four times, yielding a total of 14 occurrences for each noun.

The orders of *functor+N* versus *kes+N* trials were counter-balanced across infants. The orders of the Test trials were relative to the final Familiarization trial, with the *preuves* trial being presented first if the final Familiarization trial presented utterances containing *sangles*, and vice versa. Infants' looking times during the Test trials were measured to assess the effect of functors on the segmentation of nouns.

Procedure

The experiment was run in a sound attenuated booth. The infant sat on the parent's lap, facing a central monitor. The parent wore noise cancellation headphones that delivered masking music. Each trial, which presented a black-and-white checkerboard and the auditory stimuli simultaneously from the central position, was initiated by the infant, i.e. a researcher outside the booth pressed a computer key to start a trial when the infant looked towards the screen. Once a trial started, it continued for the entire trial length. The researcher, who was blind to the stimuli and the order of trials, pressed a computer key whenever the infant delivered a look. All looks were recorded by the computer, and the total looking time for each trial was calculated automatically. Between trials, a red light flashed on the screen to attract the infant's attention. To acquaint the infant with the procedure, we began each experimental session with a pre-trial presenting the checkerboard and sine wave analogues of repeated productions of a non-word *neem* by a different female speaker (Vouloumanos, Kiehl, Werker & Liddle, 2001). The sound was distinct from the speech stimuli and would not interfere with the experiment.

Results and discussion

The two types of sequences, *functor+N* versus *kes+N*, were each presented to infants for 45 seconds during the Familiarization phase. Thus, infants received the same amount of exposure to both types. Furthermore, infants' overall looking times to the functor sequences and the nonsense functor *kes* sequences during familiarization were equivalent, $t(31) = -.649$; $p = .521$ (two-tailed); thus, infants received comparable attentional exposure to the two types of familiarization sequences.

Test phase

Infants' looking times during the Test phase while listening to isolated nouns *preuves* and *sangles* were analyzed in a 2 × 2 mixed ANOVA, with Familiarization Type (functor context versus *kes* context) as the within-subject factor, and Functor Frequency (high versus low) as the between-subject factor. If infants used functors to segment novel words, then their listening time during the Test phase to the isolated noun which had been familiarized with a functor should be different from their listening time to the noun which had co-occurred with a nonsense syllable. Furthermore, we expected a segmentation advantage for the noun familiarized in the context of the high-frequency functor *des*, but not in the context of the low-frequency functor *mes*.

The results of the Test phase showed a significant main effect of Familiarization Type ($F(1, 30) = 5.541$, $p = .025$), but no significant effect of Functor Frequency ($F(1, 30) = 0.788$, $p = .382$), and no significant interaction of Familiarization Type × Functor Frequency ($F(1, 30) = 0.183$, $p = .672$). Infants listened significantly longer to the noun stripped from the context of a functor than to the noun from the nonsense *kes* context ((*des/mes*) Noun: Mean = 11.35 sec, *SE* = .68 sec; (*kes*) Noun: Mean = 9.79 sec, *SE* = .76 sec). These results suggest that infants at 8 months of age use functors to segment potential new vocabulary items. Infants in the experiment recognized both functors *des* and *mes*, and were not fooled by the nonsense *kes* that was matched to the real functors except for the initial consonant (see Figure 1). Recall that infants received equivalent exposure to the NPs and the *kes* sequences during familiarization. Therefore, the segmentation effect observed during the Test phase reflects what they had already learned about the functors from the natural parental input prior to the experiment. Infants used the familiar functors to interpret an adjacent syllable as a separate word.

The lack of interaction between Familiarization Type and Functor Frequency was surprising. The low frequency

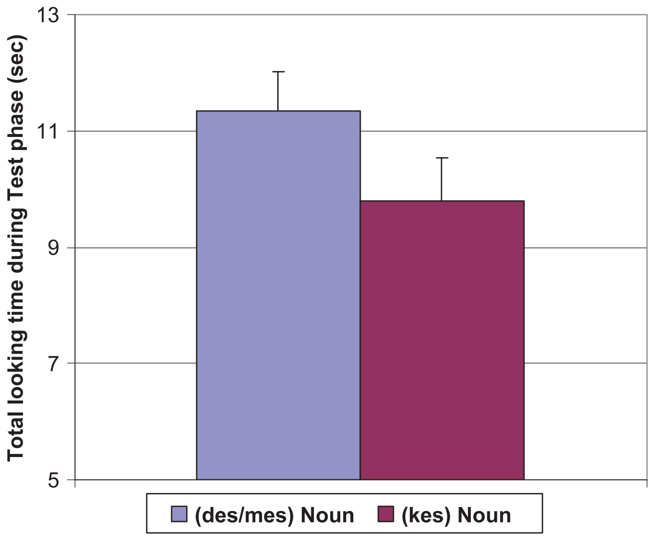

Figure 1 *Total looking time during the Test phase (with* SE*). Eight-month-old infants listened significantly longer to the isolated productions of the noun that had been previously familiarized with a frequent functor* des *or* mes*, than to the isolated noun previously familiarized with the nonsense functor* kes.

functor *mes* assisted word segmentation as did the highly frequent *des*. One possible explanation is that we may have underestimated the frequency of *mes*. In our initial analysis, only the *des* and *mes* frequencies were included, without including any homophones. However, infants at 8 months of age are unlikely to have word meanings, except for a few highly familiar content words (Fenson, Dale, Reznick, Bates, Thal & Pethick, 1994). Homophones are not likely to be distinguished at this age. The functor *mes* has a functor homophone *mais* (meaning 'but'), a conjunction that occurs commonly in the sentence initial position or in simple phrases (e.g. *mais oui*, *mais non*, etc.). We calculated the combined frequency of homophones for *des* and *mes* in the same databases. This new count showed that the frequencies were 20 for '*des*' and 22 for '*mes*' in the CHILDES transcripts, and 8802 for '*des*' versus 8193 for '*mes*' in the Beauchemin database. The two functors were thus comparable in frequency. Therefore, the significant main effect of Familiarization Type and the lack of an interaction between Familiarization Type and Functor Frequency reflect the fact that both functors are frequent from the perspective of 8-month-olds, and that both assisted word segmentation in the same fashion.

Experiment 2

To further examine the factor of functor frequency in word segmentation, we designed another experiment. If frequency is a relevant factor for infants' use of functors to segment novel words, an infrequent functor should not yield a segmentation advantage for the adjacent word.

Method

Participants

Sixteen 8-month-old monolingual Quebec-French-learning infants completed this experiment (seven males, nine females). Another five infants were tested, but their data were excluded due to fussiness (two), parental interference (one), and equipment failure (two).

Stimuli

Noun phrases were constructed with an infrequent functor *vos* (/vo/ 'your', plural form) and the same nouns used in Experiment 1. The combined frequency of *vos* and its homophones was 71 in the Beauchemin database and 0 in the two infant-directed transcripts from CHILDES used for Experiment 1. A prosodically matched functor *kos*, /ko/, differing from *vos* only in the onset consonant, was paired with the same nouns. The final stimuli set contained four tokens of each of the following sequences: *vos preuves*, *vos sangles*, *kos preuves*, *kos sangles*, as well as four tokens of *preuves* and four tokens of *sangles* in isolation. The recording, selection and manipulation of the stimuli were identical to those of Experiment 1.

Design and procedure

These were identical to those of Experiment 1. Half of the infants were familiarized with *vos preuves* trials and *kos sangles* trials, and the other half with *vos sangles* trials and *kos preuves* trials. All infants heard *preuves* and *sangles* trials during the Test phase.

Results

As in Experiment 1, *vos*+*N* and *kos*+*N* sequences were each presented for 45 seconds during the Familiarization phase. Therefore, infants received the same amount of exposure to the two types of sequences. Furthermore, infants' listening times to the *vos* NPs and the *kos* sequences during familiarization were not different ($t(15) = 1.12$, $p = .28$ (two-tailed)); thus infants' attentional exposure to the two types of sequences was equivalent.

Test phase

The data of Experiment 2 were combined with those of Experiment 1 in a 2×2 mixed ANOVA, with Familiarization Type (functor context versus nonsense context) as the within-subject factor, and Functor Frequency (high versus low) as the between-subject factor (i.e. high for Experiment 1 and low for Experiment 2). As predicted, the results showed a significant interaction of Familiarization Type × Functor Frequency ($F(1, 46) = 6.234$; $p = .016$), but no effect of Familiarization Type ($F(1, 46) = .036$; $p = .85$), nor an effect of Functor Frequency ($F(1, 46) = .153$; $p = .698$). Infants in the high-frequency

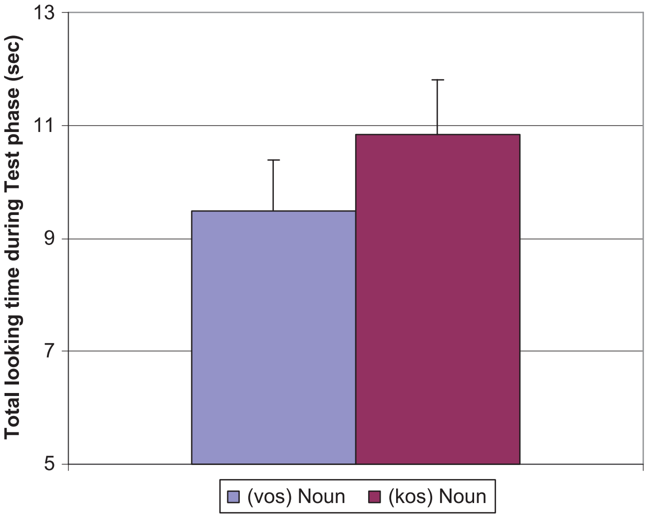

Figure 2 *Total looking time during the Test phase (with* SE*). Eight-month-old infants' listening times to the isolated productions of the noun that had been previously familiarized with an infrequent functor* vos *and to the isolated noun previously familiarized with the nonsense functor* kos*. No significant difference was found.*

condition (i.e. Experiment 1) listened significantly longer to the Test noun previously familiarized in the functor context than to the noun famliarized in the nonsense syllable context ($t(31) = 2.386$; $p = .023$, two-tailed; see details in Experiment 1). Infants in the low-frequency condition (i.e. Experiment 2) showed no significant difference in looking time while listening to the Test noun previously familiarized with the infrequent functor *vos* versus the noun familiarized with the nonsense *kos* ($t(15) = -1.347$; $p = .198$ (two-tailed); (*vos*) Noun: Mean = 9.48 sec, $SE = .91$ sec; (*kos*) Noun: Mean = 10.83 sec, $SE = .97$ sec). Thus, unlike the frequent functors *des* and *mes*, the infrequent functor *vos* produced no assistance for the segmentation of the adjacent word. The functor *vos* was perceived to be equally novel as the nonsense *kos* (see Figure 2).

General discussion

In two experiments we showed that 8-month-old French-learning infants use functional morphemes to segment potential vocabulary items from continuous speech. This indicates that functional morphemes can indeed assist infants in the earliest step of lexical learning. Moreover, frequency was shown to be the determinant for this effect. Frequent functors *des* and *mes* facilitated the segmentation of an adjacent noun. The infrequent functor *vos* did not exert this effect. Infants appeared to have processed *vos* no differently than the nonsense syllable *kos*. These findings are consistent with the existing evidence that preverbal English-learning infants use the highly frequent functor *the* to segment adjacent nouns (Shi *et al.*, 2006a). The present study demonstrates the role of functors for word segmentation in a rhythmically different language. This suggests that the effect of functors is likely universal.

Our results also suggest that infants recognize and represent specific functors *des* and *mes* and distinguish them from the prosodically matched nonsense functor *kes*, which involves only a mispronunciation at the onset position. This ability is also present in English-learning infants, where *the*, as opposed to *ke*, yields a segmentation effect for the adjacent word (Shi *et al.*, 2006a). The specific representations are, however, a few months later in English-learning than in French-learning infants, which is likely due to the differences in the prosody and distribution of functors in the two languages. In English, a stress-timing language, functors typically occur as unstressed syllables. In French, a syllable-timing language, function words are realized as more salient syllables than are functors in English. This may allow French infants to better encode specific forms. Furthermore, function words head NPs more consistently in French (e.g. *le chien*; *mon chien*; *des chiens*). English NPs do not always have a functor before a noun (e.g. *the dog*; *my dogs*; *dogs*; *big dogs*; *daycare is fun*). The French equivalent of the utterance '*Daycare is fun*' requires a determiner, '*La garderie est amusante*', not '*Garderie est amusante*'. The consistent distribution of functors in French is particularly salient and may help infants focus on functors and their patterning with the adjacent content words. Our results reveal that by 8 months of age French-learning infants not only recognize and store some specific function words, but also use them to segment other words, in preparation for vocabulary acquisition.

Our results further show that frequency is the driving factor for the segmentation effect – the functors must be frequent enough to yield such an effect. This is consistent with the effect of frequent markers shown in artificial language learning experiments with adults (Valian & Coulson, 1988). In that study adults only succeeded in learning the grammar of the language which used highly frequent markers. Mintz showed that input speech contains frequent frames distinguishing basic syntactic categories (Mintz, 2003). These frames consist largely of function words. Across natural languages the asymmetry of frequencies of occurrences for content and function words is generally observable. In comparison with content words, function words are small in the total number of types, but most functors occur highly frequently. In the Kucera and Francis (1967) database, the top 50 most frequent word types, comprising approximately half of all tokens, are nearly all function words. The same is true for the Beauchemin *et al.* (1992) Quebec-French database. Analyses of parental speech in typologically distinct languages showed the same frequency asymmetry pattern between content and function words (Shi *et al.*, 1998). Preverbal infants, as we show in the present study, are sensitive to the forms of high-frequency functors and use them to segment other novel word forms.

The early processing of function words has implications for language acquisition research. In classic studies (e.g. Brown, 1973) it was concluded that function words were learned much later than content words, in light of the observation that children's early speech production was telegraphic, missing functional elements. According to the theory of semantic bootstrapping (Pinker, 1984) and non-innatist approaches to syntactic category learning (e.g. Pine & Martindale, 1996), infants break into the grammar by acquiring some lexical words (i.e. content words) first, and only later begin to analyze functional items. There is empirical evidence that initial word–meaning association involves lexical words, particularly nouns (Gentner, 1982), rather than function words. In recent theories, however, function items are considered important for bootstrapping the earliest tasks of language acquisition (Christophe, Guasti, Nespor, Dupoux & van Ooyen, 1997; Christophe, Millotte, Bernal & Lidz, in press; Morgan, Shi & Allopenna, 1996; Shi, 1996, 2005). According to this view, infants build up rudimentary representations of the syntactic and lexical structures using functional items and prosodic cues; these representations may then bootstrap infants further into the learning of other knowledge (e.g. the role of syntax for the inference of verb meanings; Gillette, Gleitman, Gleitman & Lederer, 1999; older infants' learning of noun meanings in function word contexts; Zangl & Fernald, 2007). The results of the present study support this position: Infants shortly after the first half year of life already have access to functors and use them to segment potential vocabulary items. Moreover, this process is formal in that no processing of meaning is necessary.

In sum, this study examined how infants begin the task of extracting potential words from input speech. We suggest that it may be universal that functional elements, which occur frequently in input across languages, are not only present in initial language processing, but also help bootstrap infants into other aspects of language acquisition such as the learning of word meanings, grammatical category assignment, and phrase structure analyses.

Appendix

Average acoustic values (and standard deviations) across multiple tokens of function words during familiarization

	Vowel duration (ms)	Mean pitch (Hz)	Mean amplitude (db)
des	123.25 (26.95)	166.38 (14.8)	66.38 (2.77)
mes	128.75 (22.56)	167.13 (12.04)	65.88 (3.72)
kes	125.38 (20.13)	165.38 (11.51)	65.75 (3.81)
vos	140.38 (42.1)	195.25 (16.88)	68.88 (2.75)
kos	128 (30.66)	207.13 (5.36)	73.63 (1.92)

Acknowledgements

We thank Les Cohen for providing us with the experimental software (Cohen, Atkinson & Chaput, 2000). This research was supported by funding from NSERC, FQRSC and SSHRC. We thank Bruno Gauthier and David-Étienne Bouchard for research assistance. Partial reports of this study were presented at the XVth Biennial International Conference on Infant Studies and the Xth International Congress for the Study of Child Language.

References

Bassano, D., & Maillochon, I. (1994). Early grammatical and prosodic marking of utterance modality in French: a longitudinal case study. *Journal of Child Language*, **21**, 649–675.

Beauchemin, N., Martel, P., & Théoret, M. (1992). *Dictionnaire de fréquence des mots du francais parlé au Québec: Fréquence, dispersion, usage, écart réduit*. New York: Peter Lang Publishing.

Bernal, S., Lidz, J., Millotte, S., & Christophe, A. (2007). Syntax constrains the acquisition of verb meaning. *Language Learning and Development*, **3**, 325–341.

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.

Christophe, A., Guasti, M.T., Nespor, M., Dupoux, E., & van Ooyen, B. (1997). Reflections on prosodic bootstrapping: its role for lexical and syntactic acquisition. *Language and Cognitive Processes*, **12**, 585–612.

Christophe, A., Millotte, S., Bernal, S., & Lidz, J. (in press). Bootstrapping lexical and syntactic acquisition. *Language and Speech.*

Cohen, L.B., Atkinson, D.J., & Chaput, H.H. (2000). *Habit 2000: A new program for testing infant perception and cognition.* [Computer software]. Austin, the University of Texas.

Curtin, S., Mintz, T.H., & Byrd, D. (2001). Coarticulatory cues enhance infants' recognition of syllable sequences in speech. In A.H.J. Do, L. Dominguez, & A. Johansen (Eds.), *Proceedings of the 25th Annual Boston University Conference on Language Development* (pp. 190–201). Somerville, MA: Cascadilla.

Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., Thal, D.J., & Pethick, S.J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, **59** (Serial No. 242, No. 5).

Fernald, A., & Hurtado, N. (2006). Names in frames: infants interpret words in frames faster than words in isolation. *Developmental Science*, **9**, F33–F40.

Gentner, D. (1982). Why nouns are learned before verbs: linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language development: Vol. 2. Language, thought and culture* (pp. 301–334). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gerken, L.A., Landau, B., & Remez, R. (1990). Function morphemes in young children's speech perception and production. *Developmental Psychology*, **27**, 204–216.

Gerken, L., & McIntosh, B. (1993). Interplay of function morphemes and prosody in early language. *Developmental Psychology*, **29**, 448–457.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, **73**, 165–176.

Graf Estes, K., Evans, J.L., Alibali, M.W., & Saffran, J.R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, **18** (3), 254–260.

Hallé, P., Durand, C., & de Boysson-Bardies, B. (in press). Do 11-month-old French infants process articles? *Language and Speech.*

Höhle, B., & Weissenborn, J. (2000). The origins of syntactic knowledge: recognition of determiners in one year old German children. In *Proceedings of the 24th Annual Boston Conference on Language Development* (pp. 418–429). Somerville, MA: Cascadilla Press.

Höhle, B., & Weissenborn, J. (2003). German learning infants' ability to detect unstressed closed class elements in continuous speech. *Developmental Science*, **6**, 122–127.

Houston, D.M., Jusczyk, P.W., Kuijpers, C., Coolen, R., & Cutler, A. (2000). Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin and Review*, **7**, 504–509.

Johnson, E.K. (2005). Grammatical gender and early word recognition in Dutch. In A. Brugos, M.R. Clark-Cotton, & S. Ha (Eds.), *Proceedings of the 29th annual Boston University Conference on Language Development* (Vol. 1, pp. 320–330). Sommerville, MA: Cascadilla Press.

Johnson, E.K., & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: when speech cues count more than statistics. *Journal of Memory and Language*, **44** (4), 548–567.

Jusczyk, P., & Aslin, R. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, **29**, 1–23.

Jusczyk, P., Houston, D., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, **39**, 159–207.

Kedar, Y., Casasola, M., & Lust, B. (2006). Getting there faster: 18- and 24-month-old infants' use of function words to determine reference. *Child Development*, **77**, 325–338.

Kooijman, V.M., Hagoort, P., & Cutler, A. (2004). Word segmentation from continuous speech: an ERP study with 10-month-old infants. *Journal of Cognitive Neuroscience, Supplement*, **16**, 57.

Kucera, H., & Francis, W.N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

MacWhinney, B. (2000). *The CHILDES Project* (3rd edn.). *Vol. 2: Tools for analyzing talk: The database*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mattys, S.L., & Jusczyk, P.W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, **78**, 91–121.

Mattys, S.L., Jusczyk, P.W., Luce, P.A., & Morgan, J.L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, **38**, 465–494.

Mintz, T.H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, **90**, 91–117.

Morgan, J., Shi, R., & Allopenna, P. (1996). Perceptual bases of rudimentary grammatical categories. In J. Morgan & K. Demuth (Eds.), *Signal to syntax* (pp. 263–283). Hillsdale, NJ: Lawrence Erlbaum Associates.

Nazzi, T., Iakimova, G., Bertoncini, J., Frédonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, **54**, 283–299.

Pine, J.M., & Martindale, H. (1996). Syntactic categories in the speech of young children: the case of the determiner. *Journal of Child Language*, **23**, 369–395.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Polka, L., & Sundara, M. (2003). Word segmentation in monolingual and bilingual infant learners of English and French. In *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 1021–1024). Barcelona, Spain.

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

Shady, M. (1996). Infants' sensitivity to function morphemes. Doctoral dissertation, the State University of New York, University at Buffalo.

Shafer, V., Shucard, D., Shucard, J., & Gerken, L. (1998). An electrophysiological study of infants' sensitivity to the sound patterns of English speech. *Journal of Speech, Language, and Hearing Research*, **41**, 874–886.

Shi, R. (1996). Perceptual correlates of content words and function words in early language input (PhD dissertation, Brown University, 1995). *Dissertation Abstracts International*, **56**, 3108A.

Shi, R. (2005). Early syntactic categories in infants. In Henri Cohen & Claire Lefebvre (Eds.), *Handbook of categorisation in cognitive science* (pp. 481–495). New York: Elsevier.

Shi, R., Cutler, A., Werker, J., & Cruickshank, M. (2006a). Frequency and form as determinants of functor sensitivity in English-acquiring infants. *Journal of the Acoustical Society of America*, **119**, EL61–EL66.

Shi, R., Marquis, A., & Gauthier, B. (2006c). Segmentation and representation of function words in preverbal French-learning infants. In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *BUCLD 30: Proceedings of the 30th Boston University Conference on Language Development, Vol. 2* (pp. 549–560). Somerville, MA: Cascadilla Press.

Shi, R., Morgan, J.L., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: a cross-linguistic perspective. *Journal of Child Language*, **25**, 169–201.

Shi, R., Werker, J., & Cutler, A. (2006b). Recognition and representation of function words in English-learning infants. *Infancy*, **10**, 187–198.

Shi, R., Werker, J., & Morgan, J. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, **72**, B11–B21.

Valian, V.V., & Coulson, S. (1988). Anchor points in language learning: the role of marker frequency. *Journal of Memory and Language*, **27**, 71–86.

van de Weijer, J. (1998). Language input for word discovery. Doctoral dissertation. Max Planck Institute, Nijmegen.

van Heugten, M., & Shi, R. (in press). French-learning toddlers process gender information on determiners in online word comprehension. *Developmental Science*.

Vouloumanos, A., Kiehl, K., Werker, J.F., & Liddle, P. (2001). Neurological bases of speech and non-speech processing: an event-related functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, **13**, 994–1005.

Woodward, Z.N., & Aslin, R.N. (1990). Segmentation cues in maternal speech to infants. Poster presented at the International Conference on Infancy Studies, Montreal, Canada.

Zangl, R., & Fernald, A. (2007). Increasing flexibility in children's online processing of grammatical and nonce determiners in fluent speech. *Language Learning and Development*, **3** (3), 199–231.

Received: 4 April 2007
Accepted: 13 May 2007