Brief article

Initial morphological learning in preverbal infants

Alexandra Marquis, Rushen Shi *

Département de psychologie, Université du Québec à Montréal, C.P. 8888, Succursale Centre-Ville, Montréal, Québec, Canada H3C 3P8

ARTICLE INFO

ABSTRACT

How do children learn the internal structure of inflected words? We hypothesized that bound functional morphemes begin to be encoded at the preverbal stage, driven by their frequent occurrence with highly variable roots, and that infants in turn use these morphemes to interpret other words with the same inflections. Using a preferential looking procedure, we showed that French-learning 11-month-olds encoded the frequent French functor /e/, and perceived bare roots and their inflected variants as related forms. In another experiment an added training phase presented an artificial suffix co-occurring with many pseudo-roots. Infants learned the new suffix and used it to interpret novel affixed words that never occurred during the training. These findings demonstrate that initial learning of sub-lexical functors and morphological alternations is frequency-based, without relying on word meaning.

1. Introduction

In natural languages words may contain sub-lexical units of the root and inflectional morpheme(s). Inflectional morphemes (e.g., –*ing*) signal grammatical properties of the word without changing its core meaning. These morphemes are primarily structural, whereas the root denotes semantic properties. For example, *walking* and *walked* share the core meaning and belong to the same morphological paradigm, with –*ing* and –*ed* conveying grammatical information. How do children acquire these properties? Do they need to first learn word meaning before analyzing the internal structure of words in a morphological paradigm? We propose that morphological learning begins during infancy with the encoding of the forms of inflectional morphemes based on frequency.

Specifically, we hypothesized that prior to learning word meaning, infants can already encode the forms of roots and frequent bound morphemes separately, based on hearing variable roots co-occurring with a small set of frequent bound functional morphemes, and that when a set of bound functional morphemes are established, words such as *walk, walking* and *walked* are perceived as related because of the shared root.

The existing literature suggests that these hypotheses are plausible. Infants start perceiving functional morphemes at an early age. They distinguish forms of function words (free morphemes) from content words and encode specific function words during the first year of life (e.g., Hallé, Durand, & de Boysson-Bardies, 2008; Höhle & Weissenborn, 2003; Shafer, Shucard, Shucard, & Gerken, 1998; Shi & Werker, 2001, 2003; Shi, Werker, & Cutler, 2006; Shi, Werker, & Morgan, 1999). Little is known about the early processing of bound functional morphemes, which have the same frequency characteristics as free functional morphemes. In recent studies (Höhle, Schmitz, Santelmann, & Weissenborn, 2006; Santelmann & Jusczyk, 1998) English- and German-learning infants around 1.5 years of age preferred sentences containing non-adjacent functional morphemes that were in grammatical co-dependent relations (e.g., *is walking*) over ungrammatical relations (e.g., *can walking*), suggesting that infants must have encoded the morphemes. Mintz (2004) found that English-learning 15-month-olds perceive the suffix –*ing* and the root as separate units. Overall, these studies indicate that function words/morphemes begin to be perceived and encoded early in development.

* Corresponding author.
  E-mail address: shi.rushen@uqam.ca (R. Shi).

There are studies that offer clues to infants' knowledge of morphological alternations. Children's spontaneous speech at the early stage is typically telegraphic (e.g., *book, doggy move*), missing functional morphemes (e.g., Brown, 1973). This characteristic could indicate that children fail to perceive or encode functional morphemes, since functional morphemes are weaker in spoken forms than content words (e.g., Cutler, 1993; Monaghan, Christiansen, & Chater, 2007; Shi, Morgan, & Allopenna, 1998). However, it is possible that infants have encoded the bound morphemes but only selectively retained the meaningful root forms in their speech. There is independent evidence that the latter interpretation is likely. In a recent word learning study (Hochmann, Endress, & Mehler, 2010) Italian 17-month-olds were exposed to sentences of a foreign language, which contained high-frequency function words and low-frequency non-function words. Following the exposure, infants mapped new meaning to non-function words, but not to function words. Thus, infants expected function words to be non-semantic. In an attempt to understand the nature of omissions in children's production, Gerken, Landau, and Remez (1990) conducted an imitation study with 2-year-olds, and found that children omitted functional morphemes more than prosodically matched nonsense syllables. For example, they omitted *–es* in *Pete bounces the ball* more than they omitted *–a* in *Pete pusha ko truck*. That is, children intentionally dropped the non-semantic functional morphemes, suggesting that they may have treated the bare root and its inflected form (*bounce, bounces*) as morphological variants of the same paradigm.

The present study examined whether preverbal infants perceive morphological variants as related forms, and whether they rely on frequency information to encode the forms of bound functional morphemes. We used French as the test case. Many verbs in French contain the frequent suffix /e/ (homophonous for the infinitive *–er* and past participle *-é*), e.g., *toucher, marché, regardé*, etc. We favored using a verb functional morpheme for testing purely frequency-based encoding because verbs are generally absent in the early learning of word meaning (e.g., Gentner, 1982). Nevertheless, verb non-inflected forms (bare roots) are segmented by French-learning infants very early, at 11 months of age (Marquis & Shi, 2008). To assess infants' generalized knowledge about morphological variation, we used pseudo-roots rather than familiar roots. This ensured that infants had never heard these forms previously, thus avoiding the possibility that they may have encoded and understood the meaning of the root and the inflected versions of familiar words both as unanalyzed wholes (e.g., *touche, toucher*) based on prior experience. Moreover, in Experiment 2 an artificial training phase was introduced, presenting many pseudo-verbs containing different novel roots all co-occurring with the same artificial ending. We then tested whether this training would lead to the learning of the nonsense ending as a functional morpheme.

2. Experiment 1

2.1. Method

2.1.1. Participants

Participants were 32 monolingual French-learning 11-month-olds (mean: 345 days, SD: 6.41 days, range:

329–354 days). Two additional infants were excluded from data analysis because of fussiness.

2.1.2. Stimuli

We created two pseudo-verb roots (/trid/, /glyt/) as the target words and seven sentences for each target. Each sentence contained a target word inflected with the French verb suffix /e/, e.g., *trider-*/tride/, *gluter-*/glyte/. We also constructed seven other sentences for each target, where the targets appeared with a nonsense ending *–ou* /u/[1] (i.e. *tridou-*/tridu/, *glutou-*/glytu/). Table 1 shows the stimuli.

A female native French speaker recorded the stimuli. To obtain interesting stimuli that would help keep infants' interest in our experiment, we instructed our speaker to produce the stimuli as if she were speaking to a baby. The recordings were made in an IAC acoustic room (sampling rate 44.1 KHz, bit rate 16 bits). The final stimuli included 28 exemplars for each target (/trid/, /glyt/), and four exemplars for each sentence.

The visual stimuli included an abstract green-colored image, which was presented simultaneously with the auditory stimuli during each trial. A flashing light, synchronized with a cricket sound, served as the attention getter between trials. Water bubble sounds were used for a pre-experimental trial and a post-experimental trial together with the abstract image. The pre-trial introduced the procedure to the infant. The post-trial indicated whether infants remained on task throughout the experiment, since looking time should recover during this post-trial.[2]

2.1.3. Design

Half of the infants heard /trid/ during the familiarization phase, and the other half heard /glyt/. The tokens of the target word were presented randomly. Each familiarization trial had a maximum length of 8.5 s. The inter-stimulus interval (ISI) was 500 ms.

The test phase contained two conditions. Infants in Condition 1 all heard trials containing /trid/, versus trials containing /glyt/. The two words were inflected with the frequent suffix /e/ (e.g., *tridez-*/tride/, *gluté-*/glyte/) in sentences. Infants in Condition 2 all heard trials presenting /trid/ and /glyt/ occurring with the nonsense suffix /u/ (*tridou-*/tridu/, *glutou-*/glytu/) in sentences. The maximum

---

[1] Note that *-ou* is the spelling for /u/ in French. We chose /u/ based on a few considerations. First, it is not a bound morpheme in French, thus contrasting with /e/. Secondly, /u/ is one of the point vowels (/i/, /a/, /u/). In phonetics, point vowels are considered as more fundamental and largely language universal. Thus, we reasoned that it should be salient to infants on this ground. Thirdly, although /u/ is not a bound morpheme in French, it is a frequently occurring vowel in many French words. We conducted a frequency count of 15 vowels in the speech addressed to 8-month-old French-learning infants (from the corpus of Cécyre & Shi, 2005), and found that /u/ is among the top five most common vowels, even ranked before the /e/ vowel. The frequency count for these vowels (on a total of 3675 word tokens) from the transcripts of three mothers ranged from 948 for the most frequent vowel and 20 for the least frequent vowel. The most common seven vowels were 948 for /a/, 611 for /E/, 563 for /i/, 512 for schwa, 363 for /u/, 309 for /y/, and 242 for /e/.

[2] After each experiment and before data analysis, looking times for the last test trial and post-trial were analyzed in a paired *t*-Test. Looking times recovered significantly during the post-trial (Experiment 1: $t(31) = -7.349$, $p = .000$, 2-tailed; Experiment 2: $t(15) = -4.343$, $p = .001$, 2-tailed), suggesting that infants were on task during the experiments.

Table 1
Stimuli of Experiment 1.

Familiarization items	Group A /trid/	Group B /glyt/
Test items Condition 1 (/tride/, /glyte/)		
	Maman a /tride/ les poèmes.	Maman a /glyte/ le magot.
	/tride/ les verbes est amusant.	/glyte/ les grappes est amusant.
	T'as /tride/ la virgule.	T'as /glyte/ la parcelle.
	C'est le jeu qu'on a /tride/.	C'est le lot qu'on a /glyte/.
	Bébé va /tride/ le graphème.	Bébé va /glyte/ les gondoles.
	La phrase, j'ai /tride/.	Les roues, j'ai /glyte/.
	/tride/le juron!	/glyte/ la coquille!
Test items Condition 2 (tridou-/tridu/, glutou-/glytu/)		
	Maman a /tridu/ les poèmes.	Maman a /glytu/ le magot.
	/tridu/ les verbes est amusant.	/glytu/ les grappes est amusant.
	T'as /tridu/ la virgule.	T'as /glytu/ la parcelle.
	C'est le jeu qu'on a /tridu/.	C'est le lot qu'on a /glytu/.
.	Bébé va /tridu/ le graphème	Bébé va /glytu/ les gondoles.
	La phrase, j'ai /tridu/.	Les roues, j'ai /glytu/.
	/tridu/ le juron!	/glytu/ la coquille!

trial length was 16.5 s for Condition 1 and 16 s for Condition 2, with ISIs of 250 ms. Each trial of the same type would present all seven sentences (e.g., /tride/ sentences, see Table 1) if the infant looked during the whole trial. The order of the seven sentences was random within each trial and between trials of the same type.

The test phase presented two distinct trial types, target trials versus control trials. The sentences in the target trials contained the target presented during familiarization. Control trials contained the word that was not presented during familiarization for that group of infants. The design was counterbalanced between infants such that the target trials for one group were control trials for the other group. This counterbalancing applied for both conditions.

2.1.4. Procedure

Infants sat on the parent's lap in front of a display monitor in an IAC acoustic room. The parent wore headphones that delivered masking music. An experimental program presented the stimuli and recorded infants' looking times. Auditory stimuli were played through loudspeakers at the same location as the monitor, which presented the visual stimuli. The experimenter, blind to all stimuli, observed the infant from another room via a closed circuit TV and pressed a computer key whenever the infant looked towards the monitor. Each trial was initiated upon the infant's look towards the monitor, and terminated when the infant looked away for at least 1 s or when the end of the trial was reached. The familiarization phase terminated when the infant accumulated 30 s of looking (i.e., listening) to the target word. Test trials then started automatically. There were seven target trials and seven control trials. The two trial types were presented in alternation.

2.1.5. Predictions

If infants encoded the frequent suffix and perceived the non-inflected root and its inflected form as morphological variants, they should discriminate target versus control test trial types in Condition 1. If processing was truly based on morphology rather than partial word-form overlap, in-

fants should show no discrimination in Condition 2, where the familiarized target (e.g., /trid/) partially overlapped with the disyllabic mono-morphemic word (e.g., /tridu/).

2.2. Results

Following the standard practice in this procedure (e.g., Vouloumanos & Werker, 2004), we removed from data analysis the first test trial of each type, as they are usually unstable. For the remaining test trials, average looking time per trial for target trials and that for control trials were calculated for each infant. We conducted a 2×2 ANOVA, with Test Trial Type (target versus control test trials) as the within-subject factor, and Condition (Condition 1 real suffix /e/, versus Condition 2 non-morphemic /u/) as the between-subject factor. Results revealed no main effect of Test Trial Type ($F(1,30) = .245$, $p = .624$), no main effect of Condition ($F(1,30) = .089$, $p = .767$). But crucially, Test Trial Type $\times$ Condition interaction was significant ($F(1,30) = 4.836$, $p = .036$). Paired t-Tests were then conducted for each condition. Condition 1 (suffix /e/) showed a significant preference ($t(15) = 3.113$, $p = .007$, 2-tailed) for target over control trials. In Condition 2 (non-morphemic /u/) looking times were not different ($t(15) = -.945$, $p = .359$, 2-tailed) for target versus control trials (see Fig. 1).
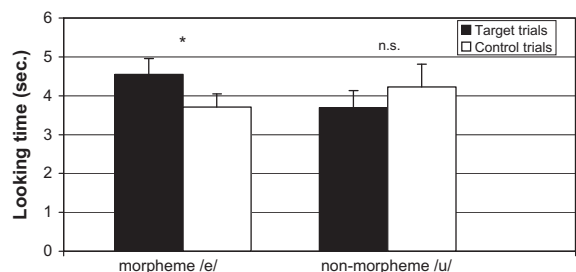
Fig. 1. Infants' looking (listening) time during the test phase to passages containing the target verb versus passages containing the control verb in conditions of real morpheme /e/ (left columns) and nonsense ending/u/ (right columns).

These results suggest that infants can encode the morpheme /e/, and can associate bare root forms with their inflected variants, e.g. /trid/-/tride/. Infants treated non-inflectional words that overlapped with part of larger words as unrelated, e.g., /trid/ as unrelated to /tridu/.

What is the mechanism underlying this initial generalized ability to parse roots and morphemes? We hypothesized that it is the highly variable roots co-occurring with frequent suffixes that enables infants to encode separately the suffixes and roots and to generalize this analysis to forms containing novel roots. In artificial language experiments, input variability appeared to be crucial for infants' generalization of rules to novel instances (e.g., Gómez & LaKusta, 2004). Here we suggest that infants had acquired from the natural input the suffix /e/ due to its frequent occurrence with many roots. The nonsense /u/, however, does not exhibit this property, and therefore was not perceived as a morpheme. In Experiment 2 we created an artificial training set in which the /u/-ending was frequent and the co-occurring roots were highly variable. We tested if infants could learn /u/ as a new suffix.

3. Experiment 2

3.1. Method

3.1.1. Participants

Participants were 16 French-learning 11-month-olds (mean: 348 days, SD: 6.91 days, range: 337–361 days). One additional infant was excluded from data analysis because of fussiness.

3.1.2. Stimuli

Stimuli were identical to those of Experiment 1, except that we added a training phase containing 14 new nonsense words sharing a common ending /u/ (Table 2). These stimuli were recorded by the same speaker as in Experiment 1. The final training set included six exemplars per word (total: $14 \times 6 = 84$). The target words in the subsequent familiarization and test phases were not among the training materials.

3.1.3. Design

All infants heard the training exemplars, presented quasi-randomly, with ISIs of 750 ms. A blue landscape image was presented simultaneously. This training, lasting 120 s, was a passive exposure phase. Training words were presented with no contingency to looking. Infants were free to do anything in the experimental room. The parent was asked not to talk to the infant. After the training, infants were presented with all the materials of Experiment 1 Condition 2. The design of the familiarization and test phases was that of Experiment 1 Condition 2.
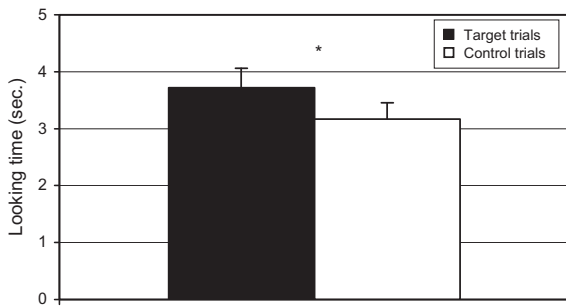
3.1.4. Procedure

Identical to Experiment 1.

3.2. Results

We compared the average looking time per trial for target test trials and control test trials, with the first trial of each type removed, as in Experiment 1. Since infants in Experiment 1 Condition 1 (morpheme /e/) showed a familiarity preference for target trials, we predicted that if the training in Experiment 2 yielded successful learning, a familiarity preference should also be observed. A one-tailed paired t-Test was therefore used. As predicted, infants preferred listening to target trials over control trials ($t(15) = 2.110$, $p = .026$, one-tailed; see Fig. 2), indicating that the training indeed enabled infants to analyze /u/ as a new suffix. Moreover, they generalized this learned knowledge in their subsequent perception of the artificial suffix and the novel roots even though the latter never occurred during training.

We also compared Experiment 2 with Experiment 1 Condition 1 (morpheme /e/) in a 2×2 ANOVA, with Test Trial Type (target versus control test trials) as the within-subject factor, and Morpheme Status (real suffix /e/, versus trained suffix /u/) as the between-subject factor. Results revealed no main effect of Morpheme Status ($F(1,30) = 2.107$, $p = .157$), no interaction of Test Trial Type × Morpheme Status ($F(1,30) = .196$, $p = .661$), but a significant effect of Test Trial Type ($F(1,30) = 13.198$, $p = .001$), indicating that Experiment 2 produced the same

Table 2
Stimuli of Experiment 2.

Pre-familiarization items
Linchou1, Cradou1, Plandou1, Wélou1, Rebou1, Balou1, Fitou1, Nangou1, Frivou1, Vaunou1, Kaurou1, Nadou1, Stakou1, Zabou1, Nangou2, Frivou2, Vaunou2, Kaurou2, Nadou2, Stakou2, Zabou2, Linchou2, Cradou2, Plandou2, Wélou2, Rebou2, Balou2, Fitou2, Vaunou3, Kaurou3, Nadou3, Stakou3, Zabou3, Linchou3, Cradou3, Plandou3, Wélou3, Rebou3, Balou3, Fitou3, Nangou3, Frivou3, Plandou4, Wélou4, Rebou4, Balou4, Fitou4, Nangou4, Frivou4, Vaunou4, Kaurou4, Nadou4, Stakou4, Zabou4, Linchou4, Cradou4, Stakou5, Zabou5, Linchou5, Cradou5, Plandou5, Wélou5, Rebou5, Balou5, Fitou5, Nangou5, Frivou5, Vaunou5, Kaurou5, Nadou5, Fitou6, Nangou6, Frivou6, Vaunou6, Kaurou6, Nadou6, Stakou6, Zabou6, Linchou6, Cradou6, Plandou6, Wélou6, Rebou6, Balou6

Familiarization items	Group A /trid/	Group B /glyt/
Test items (identical to Experiment 1 Condition 2) (*tridou-*/tridu/, *glutou-*/glytu/)		
	Maman a /tridu/ *les poèmes.*	*Maman a* /glytu/ *le magot.*
	/tridu/ *les verbes est amusant.*	/glytu/ *les grappes est amusant.*
	T'as /tridu/ *la virgule.*	*T'as* /glytu/ *la parcelle.*
	C'est le jeu qu'on a /tridu/.	*C'est le lot qu'on a* /glytu/.
	Bébé va /tridu/ *le graphème.*	*Bébé va* /glytu/ *les gondoles.*
	La phrase, j'ai /tridu/.	*Les roues, j'ai* /glytu/.
	/tridu/ *le juron!*	/glytu/ *la coquille!*

Fig. 2. Infants' looking (listening) time to target trials versus control trials during the test phase after training.

recognition response as Experiment 1 Condition 1. This analysis further supported the results of the paired *t*-Test in Experiment 2.

4. Discussion

Results of the experiments demonstrate that infants can represent frequent suffixes, and can use previously encoded functional morphemes to analyze the internal structure of newly encountered words. The results of Experiment 1 Condition 2 (non-morphemic /u/) confirm that infants did not simply perform partial phonetic mapping of the root form and its inflected version. That is, after hearing /glyt/, they did not perceive /glytu/ as containing separate units of /glyt/ and /u/. Instead, /glyt/ and /glytu/ were perceived as unrelated mono-morphemic forms. In contrast, /glyt/ and /glyte/ were perceived as related morphological variants. The inflected form (e.g., /glyte/) was perceived as containing the root (/glyt/) and the functional suffix /e/.

Experiment 2 shows that sub-lexical morphemic learning is likely driven by frequent functors co-occurring with highly variable roots. Whereas non-morphological forms (/glytu/, /tridu/) were perceived as mono-morphemic wholes in Experiment 1 Condition 2, they were encoded into sub-lexical units in Experiment 2 where the nonsense /u/ was artificially "inflected" with many different roots during training. It was striking that infants showed generalized abstract ability in interpreting the new "morpheme", since the training input did not contain the pseudo-roots used in the familiarization and test phases. Infants encoded /u/, and perceived the roots and their inflected variant (e.g., /glyt/, /glytu/) as related forms. Since both the root and suffix were nonsense forms without meaning, our results reflect form-based sub-lexical learning. Infants at the preverbal stage can initiate morphemic analysis based on frequency, without needing word meaning.

Therefore, we showed that word meaning is not required for learning morphological variants. Rather, frequency-based analysis of word-internal units can occur at an early age. We suggest that the output of this learning feeds into infants' word learning, enabling them to expect morphological variants (e.g., *walk-walking* in English, *marche-marcher* in French) to share core meaning. Our suggestion rests on the assumption that infants still interpret inflected forms as integral words (e.g., *walking, marcher*) in addition to analyzing sub-lexical units. This assumption rests on the evidence that within-word co-articulation cues are important for the perception of word-boundaries even for 8-month-olds (e.g., Curtin, Mintz, & Byrd, 2001; Johnson & Jusczyk, 2001).

The training input in Experiment 2 was designed to contain highly variable roots co-occurring with a pseudo-suffix. The transitional probability between the roots and the pseudo-suffix was set low. Infants could rely on this distributional information to analyze sub-lexical units. This mechanism would be compatible with the idea of statistical word segmentation in Saffran, Aslin, and Newport (1996). We should note that our infants may have instead used the overall high frequency of /u/ in our input without relying on the root variability. Moreover, certain other distributional properties (e.g., suffix variability) that were not included in our input may be important for sub-lexical morphemic learning. Whereas our study provides the first indication that preverbal infants can perform frequency-based morphemic analyses, the precise parameter(s) of the input distribution underlying this learning would need further experimentation.

The ability to encode bound functional morphemes can be considered in the context of the processing of function words (free morphemes). As mentioned in the Introduction, function words are perceived very early in acquisition. We demonstrate here that by 11 months infants can also process bound functors, which are more complex phonologically than free functors. The processing and representation of free and bound functional morphemes at a young age has important significance for various language acquisition tasks. By 6–8 months of age, infants begin to use function words to locate the boundary of content words (Shi, Cutler, Werker, & Cruickshank, 2006; Shi & Lepage, 2008). Grammatical categorization based on functors was observed recently in several natural languages in infants from 12 months of age (Gerken, Wilson, & Lewis, 2005; Höhle, Weissenborn, Kiefer, Schulz, & Schmitz, 2004; Mintz, 2006; Shi & Melançon, 2010). Perception of non-adjacent syntactic dependencies involving functors is present in infants close to 18 months of age (Höhle et al., 2006; Santelmann & Jusczyk, 1998; Van Heugten & Shi, 2010). Infants also use function words for online comprehension of phrases or sentences (e.g., Bernal, Lidz, Millote, & Christophe, 2007; Johnson, 2005; Lew-Williams & Fernald, 2007; Van Heugten & Shi, 2009), and in language production (e.g., Corrêa & Name, 2003; Eilers, 1975; Gerken et al., 1990). The development shown in those studies and the morphological knowledge that we show here in preverbal infants form a logical continuity of form-based syntactic acquisition from the onset of language learning.

In conclusion, we provide the first empirical evidence that infants begin strictly form-based analyses of word-internal morphology during the first year of life before learning word meaning, and that the encoding of bound functional morphemes is likely determined by the frequent occurrence of the morphemes with highly variable roots. We demonstrate that this ability is generalized and abstract, as infants can use the encoded bound morphemes

for interpreting the internal units of newly encountered words. We show that infants have rudimentary representations of morphological alternations.

Acknowledgment

References

Bernal, S., Lidz, J., Millote, S., & Christophe, A. (2007). Syntax constrains the acquisition of verb meaning. *Language Learning and Development, 3*(4), 325–341.

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

Cécyre, M.-J. & Shi, R. (2005). Speech rate in maternal speech to French-learning infants. Paper presented at the 15th meeting of the Canadian Society for Brain, Behaviour and Cognitive Science, Montreal, Canada.

Corrêa, L. M. S., & Name, M. C. L. (2003). The processing of determiner–noun agreement and the identification of the gender of nouns in the early acquisition of Portuguese. *Journal of Portuguese Linguistics, 1*, 19–43.

Curtin, S., Mintz, T. H., & Byrd, D. (2001). Coarticulatory cues enhance infants' recognition of syllable sequences in speech. In A. H.-J. Do, L. Domínguez, A. Johansen (Eds.), *Proceedings of the 25th annual Boston University Conference on Language Development* (Vol. 1, pp. 190–201). Somerville, MA: Cascadilla Press.

Cutler, A. (1993). Phonological cues to open- and closed-class words in the processing of spoken sentences. *Journal of Psycholinguistic Research, 22*(2), 109–131.

Eilers, R. E. (1975). Suprasegmental and grammatical control over telegraphic speech in young children. *Journal of Psycholinguistic Research, 4*(3), 227–239.

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

Gerken, L. A., Landau, B., & Remez, R. E. (1990). Function morphemes in young children's speech perception and production. *Developmental Psychology, 26*(2), 204–216.

Gerken, L. A., Wilson, R., & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language, 32*(2), 249–268.

Gómez, R. L., & Lakusta, L. (2004). A first step in form-based category abstraction by 12-month-old infants. *Developmental Science, 7*(5), 567–580.

Hallé, P. A., Durand, C., & de Boysson-Bardies, B. (2008). Do 11-month-old French infants process articles? *Language and Speech, 51*, 23–44.

Hochmann, J.-R., Endress, A. D., & Mehler, J. (2010). Word frequency as a cue to identify function words in infancy. *Cognition, 115*(3), 444–457.

Höhle, B., Schmitz, M., Santelmann, L. M., & Weissenborn, J. (2006). The recognition of discontinuous verbal dependencies by German 19-month-olds: Evidence for lexical and structural influences on children's early processing capacities. *Language Learning and Development, 2*(4), 277–300.

Höhle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science, 6*(2), 122–127.

Höhle, B., Weissenborn, J., Kiefer, D., Schulz, A., & Schmitz, M. (2004). Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy, 5*(3), 341–353.

Johnson, E. K. (2005). Grammatical gender and early word recognition in Dutch. In A. Brugos, M. R. Clark-Cotton, & S. Ha (Eds.), *BUCLD 29: 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.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 18*(3), 193–198.

Marquis, A., & Shi, R. (2008). Segmentation of verb forms in preverbal infants. *Journal of the Acoustical Society of America, 123*(4), EL105–EL110.

Mintz, T. H. (2004). Morphological segmentation in 15-month-old infants. In A. Brugos, L. Micciulaa, & C. E. Smith (Eds.), *Proceedings of the 28th Annual Boston University Conference on Language Development* (Vol. 2, pp. 363–374). Somerville, MA: Cascadilla Press.

Mintz, T. H. (2006). Finding the verbs: Distributional cues to categories available to young learners. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 31–63). New York: Oxford University Press.

Monaghan, P., Christiansen, M. H., & Chater, N. (2007). The phonological-distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology, 55*(4), 259–305.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old children. *Science, 274*(5294), 1926–1928.

Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for processing limitations. *Cognition, 69*(2), 105–134.

Shafer, V. L., Shucard, D. W., Shucard, J. L., & Gerken, L. A. (1998). An electrophysiological study of infants' sensitivity to the sound patterns of English speech. *Journal of Speech, Language and Hearing Research, 41*(4), 874–886.

Shi, R., Cutler, A., Werker, J. F., & Cruickshank, M. (2006). Frequency and form as determinants of functor sensitivity in English-acquiring infants. *Journal of the Acoustical Society of America, 119*(6), EL61–EL67.

Shi, R., & Lepage, M. (2008). The effect of functional morphemes on word segmentation in preverbal infants. *Developmental Science, 11*(3), 407–413.

Shi, R., & Melançon, A. (2010). Syntactic categorization in French-learning infants. *Infancy, 15*(5), 517–533.

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. F. (2001). Six-month-old infants' preference for lexical words. *Psychological Science, 12*(1), 70–75.

Shi, R., & Werker, J. F. (2003). Basis of preference for lexical words in 6-month-old infants. *Developmental Science, 6*(5), 484–488.

Shi, R., Werker, J. F., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy, 10*(2), 187–198.

Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition, 72*(2), B11–B21.

Van Heugten, M., & Shi, R. (2009). French-learning toddlers use gender information on determiners during word recognition. *Developmental Science, 12*(3), 419–425.

Van Heugten, M., & Shi, R. (2010). Infants' sensitivity to non-adjacent dependencies across phonological phrase boundaries. *Journal of the Acoustical Society of America, 128*(5), EL223–EL228.

Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: The privileged status of speech for young infants. *Developmental Science, 7*(3), 270–276.