Phrasal prosody constrains word segmentation in French 16-month-olds

SÉVERINE MILLOTTE
JAMES MORGAN
SYLVIE MARGULES
SAVITA BERNAL
MICHEL DUTAT
ANNE CHRISTOPHE

Abstract

Infants who are in the process of acquiring their mother tongue have to find a way of segmenting the continuous speech stream into word-sized units. We present an experiment showing that French 16-month-olds are able to exploit phonological phrase boundaries in order to constrain lexical access. Using the conditioned head-turning technique, we showed that infants trained to turn their head for a bisyllabic word responded more often to sentences that contained this word, than to sentences that contained both syllables of this word separated by a phonological phrase boundary. We compare these results with similar results obtained with English-speaking infants, and discuss their implication for lexical and syntactic acquisition.

Introduction

Infants who are in the process of acquiring language must learn the words of their native language, in order to build a lexicon, or mental dictionary. To do so, they must solve two complex problems: first, they have to identify and extract word forms, and second, they have to assign a meaning to each of these word forms. In this paper, we will focus on the first of these problems, and more specifically on how infants may segment the continuous speech stream in order to recover word units.

Indeed, fluent speech does not contain any obvious cues to word boundaries that would play a role equivalent to spaces in a written text. For this reason, adult listeners have been shown to rely on their knowledge of the lexicon in order to recognize words in continuous speech. At any point in time, several words compatible with the currently available phonemic information are activated, while overlapping candidates that share one or several phonemes inhibit one another. This double process of multiple activation and competition between overlapping word candidates ensures that each phoneme is ultimately assigned to one and only one word. This process has been modelled and shown to allow for efficient word segmentation (Frauenfelder & Peeters, 1990; McClelland & Elman, 1986; Norris, 1994). In addition, experimental work shows that adult listeners do rely on multiple activation and competition (e.g. McQueen, Cutler, Briscoe, & Norris, 1995; McQueen, Norris & Cutler, 1994; Norris, McQueen, & Cutler, 1995).

Even though this procedure is perfectly adequate for adult listeners, it is not appropriate for infants who do not yet possess a rich lexicon, and encounter many words that they do not know. For this reason, a number of researchers have been looking for non-lexical segmentation procedures that can work even in the absence of a lexicon. It has thus been shown that both infants and adults can rely on phonotactics, allophonic cues, and word stress, in order to achieve at least partial segmentation (see Mersad, Goyet, & Nazzi, this volume, for an up--to-date review of the infant experimental work on segmentation). In their model of word segmentation in adults, Mattys, White & Melhorn (2005) examined the relative roles of lexical cues (knowledge of the lexicon, plus semantic constraints), segmental cues (including allophonic and phonotactic regularities as well as coarticulation), and word stress cues (in English), and proposed that lexical cues pre-empted both segmental and word stress cues, at least in good listening conditions. When lexical information was damaged (either because non-words were used, or because noise was added), then adult listeners relied on segmental information and word stress cues.

We proposed that all these word segmentation procedures, whether they rely on lexical information or not, apply within the domain of prosodic units, specifically phonological phrases (Christophe, Gout, Peperkamp, & Morgan, 2003; Christophe, Guasti, Nespor, Dupoux, & van Ooyen, 1997; Christophe, Millotte, Bernal, & Lidz, 2008; Christophe, Peperkamp, Pallier, Block, & Mehler, 2004). Variations of melody, rhythm, and intensity delimit these prosodic units. Phonological phrases typically contain one or two content words together with the function words that are associated with them. They typically contain between four and seven syllables, and are characterized by preboundary lengthening (e.g. Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992) and by the fact that there is one melodic contour per phonological phrase (Hayes & Lahiri, 1991, for Bengali; Pasdeloup, 1990, for French). Newborn infants have been shown to perceive the cues that correlate with phonological phrase boundaries (Christophe, Dupoux, Bertoncini, & Mehler, 1994; Christophe, Mehler, & Sebastián-Gallés, 2001), and 9-month-old infants have been shown to react to the disruption of phonological phrases in whole sentences (Gerken, 1994; Kemler-Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989).

In previous work, we showed that phonological phrase boundaries block lexical activation (Christophe et al., 1994; Millotte, Frauenfelder, & Christophe, 2007; see also Salverda, Dahan, & McQueen, 2003; Shukla, Nespor, & Mehler, 2007). Thus, French adults who were asked to detect a monosyllabic word (such as 'chat' / *cat*) were slowed down when that word belonged to a string of syllables with a local lexical ambiguity, showing evidence of multiple activation (e.g. [un chat grincheux], *a grumpy cat* containing the potential competitor word 'chagrin' *sorrow* was processed more slowly than [un chat drogué] *a doped cat* that contains no potential competitor, since no word in French start with 'chad…'). In contrast, when the lexical competitor straddled a phonological phrase boundary, there was no delay in lexical recognition (e.g., [son grand chat] [grimpait...], *his big cat was climbing* potential competitor 'chagrin', was not delayed relative to the non--ambiguous control). These results show that a potential lexical competitor that straddles a phonological phrase boundary does not get activated: this prosodic boundary is perceived as signalling the end of the current word. Further work showed that prosodic word boundaries (minor prosodic boundaries within a phonological phrase) also influenced subjects' performance, even though they were not powerful enough to completely block the activation of straddling competitors (Millotte, Frauenfelder, & Christophe, 2007). In infants, we found that American 10- and 13-month-olds who were trained to turn their head for the word 'paper' responded much more often to sentences that actually contain the target word, as in '[The scandalous paper] [sways him] [to tell the truth]' than to sentences that contained both of its syllables separated by a phonological phrase boundary, as in '[The outstanding pay] [persuades him] [to go to France]' (Gout, Christophe, & Morgan, 2004). Thus, American infants also perceived phonological phrase boundaries and interpreted them on-line as word boundaries.

Experiment

In this experiment, we tested whether French infants also exploit phonological phrase boundaries to constrain lexical access. As in Gout, Christophe & Morgan (2004), we used a conditioned head-turn procedure to provide an on--line measure of infants' word detection. Infants participated in two experimental sessions. During an initial training session, infants learned to turn their head upon hearing a particular word. One group of infants was trained on a bisyllabic target (either 'balcon' *balcony* or 'vipère' *viper*); a second group was trained on a monosyllabic target that matched the first syllable of one of the bisyllabic targets (either 'bal' meaning *ball*, where people dance, or 'vie' *life*). During the test session, infants heard sentences containing or not these targets. Some sentences contained the bisyllabic target itself ('balcon' or 'vipère'), while others contained both its syllables separated by a phonological phrase boundary, as in the following sentences (where square brackets indicate phonological phrases):

[Le joli balcon] [cloisonnait la terrasse] 'balcon'-sentence
[The lovely balcony] [divided the terrace].
[Le dernier bal] [conclura la saison] 'bal][con'-sentence.
[The last ball] [will conclude the season].

We expected infants trained on bisyllabic targets to turn their head more often when the target did not straddle a phonological phrase boundary than when it did. Infants from the monosyllabic group were tested on the same sentences, and were expected to show the reverse pattern of results, turning more often for 'bal][con'-sentences that actually contain the target word 'bal' than for 'balcon'-sentences that contain a syllable homophonous to the target word. Note, however, that phonological phrase boundaries should be sufficient, though not necessary, for locating word boundaries. Indeed, most phonological phrases contain more than one word, so that many word boundaries do not coincide with a phonological phrase boundary. Monosyllabic targets either immediately preceded a phonological phrase boundary (when the associated bisyllable straddled the boundary, ...bal] [conclura...]) or constituted one syllable of a continuing prosodic group (...balcon...). Although the presence of a phonological phrase boundary located just after the targets might aid infants in segmenting the target from the sentences, the absence of such a boundary should not preclude them from doing so. We therefore expected a smaller difference between sentence types for infants trained on monosyllabic targets.

A pilot experiment was run with 13-month-old French infants (the same age as the older American group from Gout et al., 2004). However, we found that, in contrast with American infants, most French 13-month-olds failed to segment bisyllabic words from the middle of sentences: indeed, most of the infants trained on 'balcon' did not turn their head for 'balcon'-sentences relative to distractor sentences that did not contain any word resembling 'balcon'. This result confirms other studies suggesting that French infants become able to segment bisyllabic words from fluent speech later than American infants (Gout, 2001; Nazzi, Iakimova, Bertoncini, Frédonie, & Alcantara, 2006, as well as the discussion in Mersad et al., this volume). We will come back to this point in the general discussion. For the present study, we tested slightly older infants, of 16 months of age.

Method

Participants:

Thirty six French-learning infants, 16 months of age, participated in this two-session experiment and were split between two groups (20 in a bisyllabic group and 16 in a monosyllabic group). One hundred and four additional infants were tested but were not included in the analyses: 71 were excluded during the first session (see procedure) because they were either not interested, or afraid, or not cooperative (because of illness or fatigue), or because they did not come back for the second session although they had succeeded in the first one. The other 33 infants were rejected during the second session because they were either not interested in the reinforcement anymore, or fussed out before completing the session. Three factors account for this high drop-out rate: the experiment takes place in two separate sessions, each of these sessions is fairly long (10-15 minutes for the first session, and 20 minutes for the second one), and it is difficult to get infants of 16 months to sit quietly through an experiment.

Stimuli:

The stimuli used in the first session were isolated words: tokens of the bisyllabic words 'balcon' (*balcony*) and 'vipère' (*viper*) for the bisyllabic group, and tokens of the monosyllabic words 'bal' (*ball* where people dance) and 'vie' (*life*) for the monosyllabic group. We selected 'balcon' and 'vipère' because the first syllables are real words in French, and because many polysyllabic verbs start with the second syllables. This allowed us to construct many different test sentences as described below.

The stimuli used in the test phase were sentences. For each target word, we constructed 12 pairs of sentences such that one member of each pair contained the bisyllabic word itself while the second member contained both syllables of the word separated by a phonological phrase boundary (see Appendix for a complete list of materials).

Sentences of each pair were matched in their prosodic structures before the target word, as well as in total number of syllables and in the number of syllables before and after the target words. Sentences were read in a motherese-like manner by a female French speaker (the last author). The same speaker read the target words in isolation, as well as short sentences with the target words placed either in final position (such as 'c'est le balcon' / *it is the balcony*) or in medial position (such as 'j'aimerais que mon balcon soit tout en bois' / *I would like my balcony to be made of wood*) which were presented in a short review phase at the beginning of the second session (see procedure).

The sentences were recorded in a sound-proof room, and the stimuli were digitized at 16 Khz and 16 bits through an OROSAU22 soundboard. The beginning of the target words were manually marked using the Bliss software.

Waveforms and pitch contours of example sentences are shown on Figure 1. We analyzed the acoustic realizations of the phonological phrase boundaries by measuring the duration of each phoneme in the bisyllabic targets (see Table 1), as well as the fundamental frequency and the energy of each vowel (see Table 2) using Praat (http://www.fon.hum.uva.nl/praat/; Boersma, 2001).

The duration analyses showed different duration patterns depending on the position relative to phonological phrase boundaries. As expected from the literature (e.g. Wightman et al., 1992), we found a significant phrase-final rhyme lengthening (between 50% and 145% depending on segments): segments situated at the end of a phonological phrase were much longer than the same ones located in the middle of a word. Thus, /al/ and /i/ were longer when followed by a phonological phrase boundary (in 'bal][con'-sentences) than when they were the first syllable of a bisyllabic word (lines 2 and 3 of table 1). Similarly, /õ/ and /Er/ were longer when followed by a phonological phrase boundary (in 'balcon'-sentences) than when they belonged to the first syllable of a multisyllabic verb.

Table 1: Duration of the segments comprising the critical sequences ('balcon' and 'vipère').

		PP boundary bal] [con vie] [per	No boundary balcon vipère	Difference		t-test		Lengthening (%)
		Mean (ms)	Mean (ms)	Mean	SE	$t(23)$	p	
Onset 1	b v	102	73	29	7.1	4.1	<.001	39.8
Vowel 1	a i	306	129	177	19.8	9.0	$<10^{-9}$	137.4
Coda 1	l	285	124	161	6.2	18.4	$<10^{-9}$	129.3
Onset 2	k p	124	120	4	8.2	<1	0.7	3
Vowel 2	õ ɛ	124	303	-180	18.8	9.6	$<10^{-9}$	-144.9
Coda 2	r	88	132	-44	4.8	5.0	$<10^{-5}$	-49.5

The pitch analyses also showed different pitch contours depending on the position of the segments relative to phonological phrase boundaries. As expected from the literature on French prosody (Di Cristo, 2000; Welby, 2003, 2006), we found a significant pitch rise at the end of each phonological phrase. In 'bal][con'-sentences, a phonological phrase boundary was placed after the 'bal' syllable and we found a significant increase of 78 Hz on this 'bal' syllable (as visible in ellipse 1 in Fig. 1; the pitch difference was measured between the first vowel of the target word and the preceding one, see the line V1-V0 in table 2). In 'balcon'-sentences, the phonological phrase was placed after the second syllable of the target word and we found a significant increase of 119 Hz on the 'con' syllable (ellipse 2 on Fig. 1, see the line V2-V1 in table 2).

Finally, the energy analyses conducted on the vowels of the target syllables showed a difference according to their position: when it was the last syllable of the target word the energy was 0.104, and when it was the first syllable of the verb that followed the target word the energy was 0.148. The syllable is thus produced with a higher energy at the beginning of a word (and at the beginning of a phonological phrase) than at the end of a word (and at the end of a phonological phrase). This is congruent with the initial strengthening described in the literature (Cho & Keating, 2001; Fougeron & Keating, 1997; Keating, Cho, Fougeron, & Hsu, 2003).

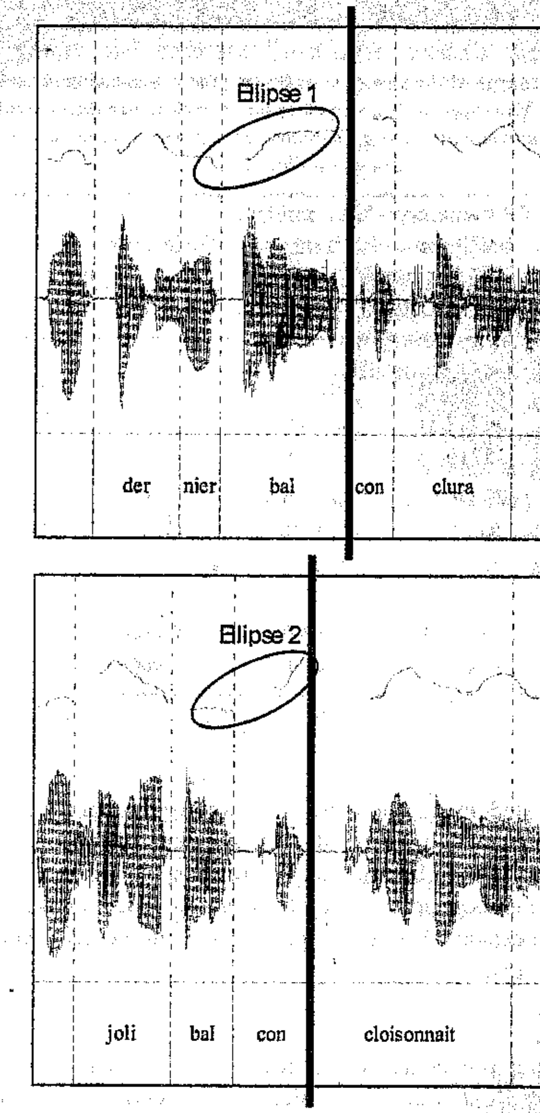

Figure 1: Waveform and fundamental frequency graph. The 'bal][con'-sentence is represented at the top and the 'balcon'-sentence at the bottom. The black bars show the position of the phonological phrase boundaries. The pitch rise at the end of a phonological phrase can be seen in ellipses 1 and 2.

To summarize, these acoustic analyses showed that sentences were realized with a very significant phrase-final lengthening as well as with a pitch rise at the end of a phonological phrase (compared to the same phonemes in word-medial position). Importantly, we also observed that none of the sentences had a pause at the phonological phrase boundary.

Table 2: Pitch and energy of the vowels constituting the critical sequences ('balcon' and 'vipere', V1 and V2 respectively), as well as one vowel before the critical sequence to compute pitch movements before phonological phrases (visible on the ellipses drawn on Fig. 1).

	PP boundary **bal] [con Vie] [per**	No boundary **balcon vipère**	Difference	t-test	
	Mean (SE)	Mean (SE)	Mean	$t(23)$	p
Pitch (Hz)					
V0	272 (8.4)	307 (11.2)	-35	2.9	<.01
V1 (a / i)	349 (8.3)	273 (8.7)	76	6.9	$<10^{-6}$
V2 (õ/ ε)	357 (9.1)	392 (5.4)	-35	3.5	<.01
V1-V0	77 (10.0)	-34 (16.5)	111	6.2	$<10^{-5}$
V2-V1	8 (12.3)	119 (10.7)	-111	7.2	$<10^{-6}$
Energy (rms)					
V1 (a / i)	0.155 (0.01)	0.154 (0.01)	0.001	<1	
V2 (õ / ε)	0.148 (0.01)	0.104 (0.01)	0.044	4.5	$<10^{-3}$

Apparatus:

Infants were tested in a sound-treated laboratory room. Trial duration, stimulus presentation, and delivery of reinforcement were controlled by a custom-designed software. Stimuli were presented through loudspeakers located in the testing room to the left of the infant. Four smoked Plexiglas boxes containing mechanical toys that provided reinforcement were located above the loudspeakers.

Procedure:

The experimental procedure was a two-session variant of Conditioned Head-Turning (as in Gout et al., 2004). Each of the sessions was divided into two main phases. In the first session (training), infants completed a shaping and a criterion phase. In the second session (maximum 7 days later), infants first went through a review phase, then completed the test phase *per se*.

Throughout both sessions, infants were seated on their parents' laps at a small table. An assistant seated directly across from the infant maintained the infant's attention at midline by silently displaying and manipulating an assortment of toys. The loudspeakers and the teddybears that provided reinforcement were located 90° from midline on the infant's left, about 1.5 meter away; a video camera was located directly above the loudspeakers.

Another experimenter in the control room observed the infant on a video monitor and judged whether the infant looked into the camera. Throughout all sessions, parent and assistant listened to acoustic masking over noise-attenuation headphones. At any time, the experimenter could also ask the assistant to change her behaviour, via a microphone connected to the assistant's headphones. The experimenter initiated trials when infants' attention was focused at midline by pressing the left mouse button. When the infant turned its head towards the loudspeakers, the experimenter pressed the right mouse button to signal a head-turn. The computer delivered reinforcement only if it was an appropriate head turn (i.e. to a target word).

The training session comprised the shaping and the criterion phases. During this session, infants heard a background word which was played continuously, presented at a comfortable listening level (68 dB SPL-b) with 1000 ms inter-stimulus interval. When a stimulus was delivered, the background word was replaced by three repetitions of the target word, allowing for a response window of 4 seconds. Infants in the bisyllabic group heard the words 'balcon' (*balcony*) and 'vipère' (*viper*) and infants in the monosyllabic group heard the words 'bal' (*ball*) and 'vie' (*life*). For each infant, one of the words served as target and the other one as background (half the infants in the bisyllabic group heard 'balcon' as target and the other half heard 'vipère' as target; similarly, half the infants in the monosyllabic group heard 'bal' as target and the other half heard 'vie' as target). Only head-turns occurring while the target word was being played were reinforced, in order to teach infants to turn their head toward the loudspeakers whenever they heard the target word. During the shaping phase, all trials were change trials. The target word was initially presented at a level 12 dB higher than that of the background word to elicit orienting head-turns. The intensity difference between target and background words was decreased in 4 dB steps each time the infant correctly responded to a change trial, until both stimuli were presented at equal intensity levels. When the infant failed to turn on three consecutive trials, the sound level was increased by 4 dB. The shaping phase continued until 30 trials were completed or until the infant turned its head on two consecutive trials with equal target and background sound levels.

At this point, the criterion phase began. This phase was similar to the shaping phase, except that trials were either change trials or no-change trials (50% of each, randomly selected by the computer). Infants were tested until they reached the predetermined criterion of 7 correct responses out of 8 consecutive trials (by turning on change trials and not turning on no-change trials). When an infant failed on three consecutive trials, retraining trials were introduced following the same schedule as during shaping, until the infant correctly turned on two consecutive trials with equal intensity. Infants not reaching criterion within 40 trials were excluded from further participation; infants who reached criterion returned for the test session within a maximum of 7 days.

The testing session began with two review phases: in the first one, the background and the target words were presented sentence-finally in short sentences (e.g. 'Regarde la vipère!' *Look at the viper!* or 'C'est le balcon!' *It is the balcony!*). In the second one, the target or background words appeared in the middle of slightly longer sentences (such as 'J'aimerais que mon balcon soit tout en bois' *I would like my balcony to be made of wood* or 'La petite vipère se prélassait au soleil' *The small viper was resting in the sun*). When a trial was requested, the infant had 3.5 seconds, starting from the beginning of the target word, to respond; the target sentence was repeated twice (in the first review phase), or once (in the second review phase). The first trials of each phase were delivered at an intensity level 8 dB higher than the background level. This intensity was progressively lowered in 4 dB steps each time the infant correctly turned its head to the loudspeakers until it reached the background intensity level.

The test phase then began. The infants were presented with the 24 experimental sentences of their group (12 'balcon'-sentences and 12 'bal][con'-sentences, or 12 'vipère'-sentences and 12 'vie][per'-sentences]. The background was constituted by 'vipère'- and 'vie][per'-sentences for the 'balcon' was constituted by group (and vice-versa) which were continuously played until a trial was requested. Infants had 2.5 seconds to respond, starting from the beginning of the test word. If an infant did not turn its head on 3 consecutive trials, the second review phase started again until the infant was successful, then the test phase resumed.

Results

For each infant included in the analyses, the testing session was recoded off-line by an experimenter who was blind to what the infants were listening to, and to the group they belonged to.

Two ANOVAs were conducted on the mean percentages of head-turns for test sentences (see Figure 2), one with participants and one with items as random factors. The by-subject analysis included two between-subjects factors: Group (bisyllabic vs. monosyllabic) and Material (counterbalancing factor: infants responding to balcon/bal vs. vipère/vie). There was also one within-subject factor: Sentence Type ('balcon'- vs 'bal][con'-sentences). The by-item analysis included one between-items factor (Material) as well as two within-item factors (Group and Sentence Type).

The ANOVAs revealed a main effect of Sentence Type in the subjects analysis only ($F_1(1,32)=4.8$, $p=.04$; $F_2(1,22)=2.1$, $p=0.2$) as well as a main effect of Material ($F_1(1,32)=5.1$, $p=.03$; $F_2(1,22)=6.2$, $p=.02$; $minF'(1,54)=2.8$, $p=0.1$). Crucially, the interaction between Group and Sentence Type was highly significant ($F_1(1,32)=45.3$, $p<.001$; $F_2(1,22)=37.9$, $p<.001$) reflecting the fact that infants from the bisyllabic group responded more often to 'balcon'- sentences than to 'bal][con'- sentences (48,2% versus 21,7%, $F_1(1,18)=32.6$, $p<.001$; $F_2(1,22)=21.8$, $p<.001$) whereas infants from the

monosyllabic group showed the reverse pattern with more headturns on 'bal][con'- sentences than on 'balcon'- sentences (27% versus 44.2%, $F_1(1,14)=15.7$, $p<.001$; $F_2(1,22)=20.7$, $p<.001$).

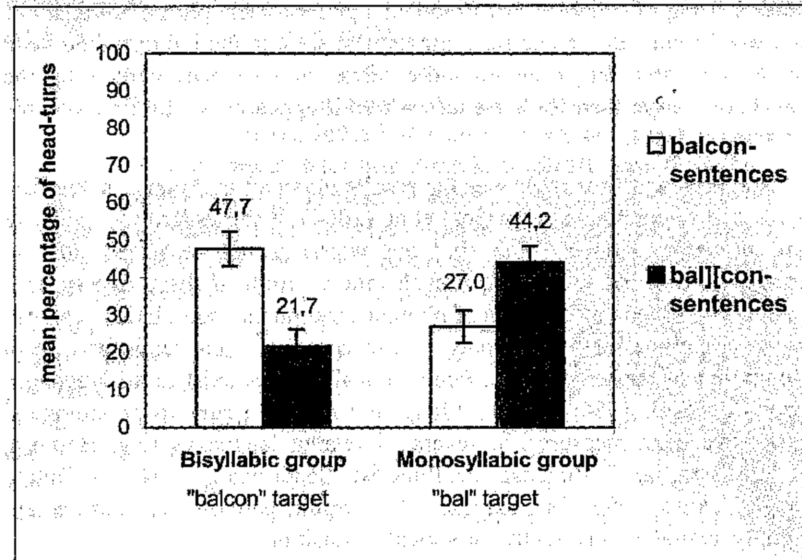

Figure 2: Mean percentage of head-turns for 16-month-old French infants in the bisyllabic group (on the left) and in the monosyllabic group (on the right). Infants heard sentences containing the bisyllabic word (balcon-sentences, white bars) and sentences containing its two syllables separated by a phonological phrase boundary (bal][con-sentences, grey bars).

We also computed an A' for each infant (non-parametric version of the d'), using the number of hits, misses, false alarms and correct rejections (Grier, 1971). We obtained a mean A' of 0.72 (standard error = .03) for the bisyllabic group, which was significantly different from chance ($t(19)=6.6$, $p<.001$). For infants in the monosyllabic group, the mean A' was 0.64 (standard error = .03) also significantly different from chance ($t(15)=4.2$, $p<.01$) but weaker than the one obtained in the bisyllabic group ($t(35)=2.09$, $p=.04$ for a direct comparison between the A' of the bisyllabic and the monosyllabic groups).

Discussion

This experiment shows than 16-month-old French infants are able to infer the presence of a word boundary when they hear a prosodic boundary. Two sequences that are phonetically identical (same segments between 'balcon' and 'bal][con') but prosodically different (presence or absence of a phonological phrase boundary) were considered as different by young infants. When they were trained to respond to a bisyllabic target, they responded more often when hearing the whole bisyllabic word within a sentence, than when they heard both its syllables separated by a prosodic boundary. They did not consider a string of syllables that straddled a phonological phrase boundary as a good word candidate. In contrast, infants who were trained to turn their head for a monosyllabic target turned more often for sentences containing the monosyllabic target than for sentences where it appeared as the first syllable of a bisyllabic word.

This experiment thus replicates the results obtained in American English, with 10- and 13-month-olds, by Gout et al. (2004). It shows that the ability to exploit prosodic boundaries to find out words is not restricted only to American infants, but extends to French infants. Both of these experiments also showed that the computation of phonological phrase boundaries is an on--line process, since infants had only 2.5s to complete their head-turns, from the beginning of the target word. Phonological phrases exist in all languages (Nespor & Vogel, 1986; Selkirk, 1984), and prosodic cues to phonological phrases have been measured in several unrelated languages (e.g. Barbosa, 2002, for Brazilian Portuguese; de Pijper & Sanderman, 1994, for Dutch; Fisher & Tokura, 1996, for Japanese). As a result, phonological phrases are potentially available universally for lexical acquisition.

In addition, we observed that French infants were able to perform the segmentation task successfully at a later age than American infants. Indeed, 10-month-old American infants already showed a differential response rate to 'paper'- vs 'pay][per'-sentences, thus exhibiting an ability to segment the words from the continuous speech stream. In contrast, our pilot study with 13--month-old French infants showed that only a small minority of them were able to correctly respond to 'balcon' in 'balcon'-sentences, relative to distractor sentences that did not contain anything remotely resembling the target word, 'balcon'. At 16 months, French infants performed significantly better than chance. There is thus between 3 and 6 months of delay between French and American infants as regards their on-line word segmentation abilities. This result is congruent with other studies, using different experimental techniques, that also showed a delay in the ability of French infants to segment bisyllabic words from fluent speech, relative to American infants. Thus, Gout (2001), using the variant of the head-turn preference procedure adapted to test word segmentation by Jusczyk & Aslin (1995), found that French infants became able to segment bisyllabic words only after the age of 12 months (at about 13 months), whereas American infants already succeed at 7 months, at least on the most frequent word type (Jusczyk & Aslin, 1995; Jusczyk, Houston, & Newsome, 1999). Similarly, Nazzi et al. (2006) found that 8-month-old French-learning infants were not able to segment bisyllabic words from fluent speech. They showed this ability at 16 months only. Before this age, French infants seemed to segment the speech

stream into syllable-sized units: at 12 months of age, they extracted the final (or initial) syllables of the target bisyllabic words rather than segmenting the target words as coherent units (the word 'toucan' would thus be represented as the syllable 'tou' followed by the syllable 'can'). Bisyllabic words were segmented as whole units later, at 16 months (see Mersad et al., this volume, for a full discussion of these results).

If it is actually the case that French is harder to segment than American English, at least for young infants, then one expects to find a delay in the development of the lexicon. Indeed, if French infants become able to segment words out of fluent speech later than American infants, then they should start building their lexicon slightly later. In addition, longitudinal data gathered by Newman et al. (2006) with American infants shows that speech segmentation ability is the best predictor of vocabulary at 2 years. To check whether French infants are delayed in their vocabulary development, we examined the vocabulary data existing in the literature, and gathered with the McArthur Communicative Development Inventory (Fenson, Dale, Reznick, Bates, Thal, & Pethick, 1994). As can be seen in Figure 3, French infants score consistently lower than American infants on the CDI, both in comprehension and in production, with a delay of about a month in the earliest stages, up to about 2-3 months later on (e.g., French 16-month-olds understand 40% of the words tested on the CDI, whereas 13.5-month-old American infants already achieve this comprehension score).

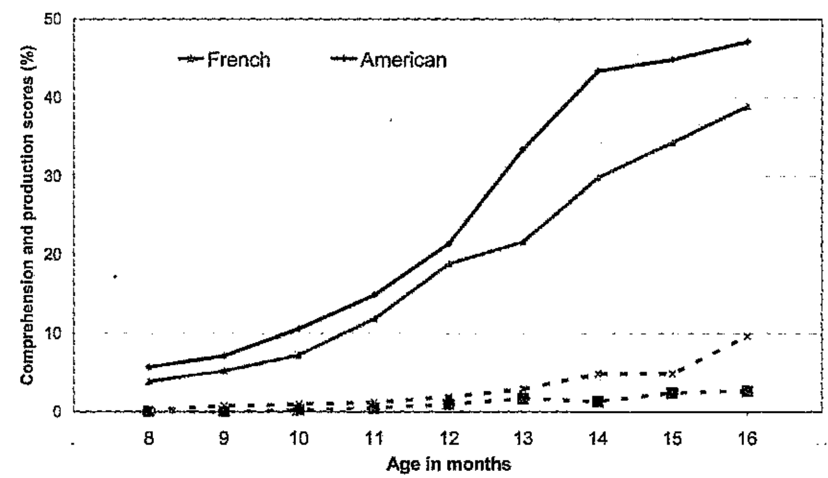

Figure 3: Comprehension and production scores, in percentages, on the American and French CDI. American data from Fenson et al. (1994), French data from Kern (2007). Comprehension scores are shown in full lines, production scores in dotted lines.

What are the factors that may account for the fact that French appears to be harder to segment than English? One factor that may play a role is the fact that in English, strong and weak syllables alternate, whereas in French, all syllables possess an approximately equal weight. It has been suggested that strong syllables may act as anchors for English listeners, who would therefore find it easy to segment bisyllabic items when the first syllable is strong (see e.g. Jusczyk et al., 1999). In contrast, French infants would start by segmenting individual syllables from the continuous speech stream, and therefore they would not initially find bisyllabic words easily (see e.g. Nazzi et al, 2006). As presented in Mersad et al. (this volume), initial segmentation strategies would thus depend on the rhythmicity of the native language: stress--based for English (easy to segment trochaic bisyllabic units) and syllable--based for French (easy to segment syllable one at a time, hard to segment polysyllabic words). A second factor that may differentially influence early segmentation abilities is the cultural differences in the way adults speak to infants. In American English, motherese is typically very exaggerated, with wide pitch excursions and a very slow speech rate. Since both speech rate and positive emotional prosody were reported to positively influence infants' ability to segment continuous speech into words (Morgan, Singh, Bortfeld, Rathbun, & White, 2001; Singh, Bortfeld, Rathbun, & Morgan, 2000; Singh, Morgan, & White, 2004), more exaggerated motherese might contribute to the better segmentation performance of American infants. As an aside, British English is informally reported to have a less exaggerated motherese as American English, and the performances of British English infants on the MacArthur CDI also appears to be lower (see Hamilton, Plunkett, & Schafer, 2000). Different styles of motherese might thus contribute to differences in segmentation abilities, even though this would need to be formally tested in cross-linguistic and cross-cultural studies. Whatever the reason why French--speaking infants become able to segment words from continuous speech later than American-speaking infants, this result stresses the necessity for cross--linguistic research. Indeed, even though one may come up with possible interpretation for this result, it was definitely unexpected (which is why several labs performed several experiments before recognizing the fact that French infants' failure to segment was a *bona fide* failure to segment, rather than a technical problem with the experimental procedure).

Apart from the fact that French infants become able to segment words from the continuous speech stream later than American infants, the two studies show exactly parallel results: infants learning both languages exploit phonological phrase boundaries in order to constrain lexical access. In addition to being helpful for lexical segmentation, phonological phrases may also be crucial for the first steps of syntactic acquisition. Indeed phonological phrases may provide some information as to the syntactic structure of sentences (Gerken, 1994; Morgan, 1986). It has often been claimed that prosodic structure may help bootstrap syntactic acquisition (Christophe, Nespor, Guasti, & van Ooyen, 2003; Gleitman & Wanner, 1982; Hirsh-Pasek, Nelson, Jusczyk, Cassidy, Druss, & Kennedy, 1987; Jusczyk, 1997). This information can be exploited in two different ways. Firstly, to constrain distributional analyses of the speech input. Thus, function words and morphemes tend to occur at the edges of syntactic phrases, and therefore also at the edges of phonological phrases. Their position within phonological phrases may thus be one of the cues that distinguish function words from content words (Morgan, Shi, & Allopenna, 1996; Shi, Morgan, & Allopenna, 1998; Shi, Werker, & Morgan, 1999). Infants could compile a list of the syllables that occur at the beginning and end of prosodic units, storing the most frequent syllables and subsequently identifying these syllables as closed--class items. Several experimental studies have shown that infants younger than 12 months already know some of the function words of their native language (see Hallé, Durand, & de Boysson-Bardies, 2008, for French; Shafer, Shucard, Shucard, & Gerken, 1998; Shi, Cutler, Werker, & Cruickshank, 2006; Shi & Gauthier, 2005; Shi, Werker, & Cutler, 2006, for English). In addition, infants in their second year of life already form categories of function words (e.g. articles that go with nouns, pronouns that go with verbs, see Höhle, Weissenborn, Kiefer, Schulz, & Schmitz, 2004; Kedar, Casasola, & Lust, 2006; Shi & Melançon, in press; Zangl & Fernald, 2007), and are even able to exploit these in order to infer the probable meaning of an accompanying content word (noun vs. verb, Bernal, Lidz, Millotte, & Christophe, 2007; Waxman, Lidz, Braun, & Lavin, 2009).

Secondly, phonological phrase boundaries may also be exploited directly by infants to constrain on-line syntactic analysis. Phonological phrase boundaries systematically coincide with boundaries of syntactic constituents (even though the reverse is not true); as a result, some syntactic constituents are prosodically marked. Thus, recent work showed that adult listeners are able to exploit phonological phrase boundaries in order to disambiguate the syntactic category of an ambiguous word (Kjelgaard & Speer, 1999; Millotte, René, Wales, & Christophe, 2008; Millotte, Wales, & Christophe, 2007; see also Name & Silva, 2009). Note that prosody in itself provides no cue to the labelling of constituents (into e.g. Noun Phrase, Verb Phrase, etc.). However, prosodic information may be used in conjunction with function words to perform an initial segmentation and labelling of sentences into syntactic constituents. Thus, the sentence "the little boy is running fast" may be initially perceived as [the xxx]NP [is xxx]VP, where the boundaries are given by prosody and the labelling is given by function words, since we know that infants build function word categories early on. Simultaneous access to both function words and phonological phrase boundaries may thus allow young children to start constructing rough syntactic analyses for the sentences they hear, even before they know the meaning of many content words (Christophe et al., 1997; Christophe et al., 2008). In fact, such a skeleton of a syntactic structure, or *syntactic skeleton*, may be what is needed to help learning the meaning of words (Gleitman, 1990).

Acknowledgments

This work was supported by a PhD and a post-doctoral fellowship from the Direction Générale de l'Armement (France), as well as a Marie Curie Intra-European Fellowship within the 6th European Community Framework Programme (n° 024843) to Séverine Millotte. This work was also supported by a grant from the French Agence Nationale pour la Recherche ('Early Language Acquisition: Experimental and Computational Approaches') to Anne Christophe as well as NIH award HD32005 to James Morgan.

References

Barbosa, P. A. (2002). Integrating gestural temporal constraints in a model of speech rhythm production. In S. Hawkins & N. NGuyen (Eds.), *Proceedings of the ISCA workshop on Temporal Integration in the Perception of Speech* (pp. 54). Cambridge: Cambridge University Printing Service.

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

Boersma, Paul (2001). Praat, a system for doing phonetics by computer. *Glot International, 5*, 341-345.

Cho, T., & Keating, P. (2001). Articulatory strengthening at the onset of prosodic domains in Korean. *Journal of Phonetics, 28*, 155-190.

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America, 95*, 1570-1580.

Christophe, A., Gout, A., Peperkamp, S., & Morgan, J. L. (2003). Discovering words in the continuous speech stream: The role of prosody. *Journal of Phonetics, 31*, 585-598.

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

Christophe, A., Mehler, J., & Sebastián-Gallés, N. (2001). Perception of prosodic boundary correlates by newborn infants. *Infancy, 2*, 385-394.

Christophe, A., Millotte, S., Bernal, S., & Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. *Language and Speech, 51*, 61-75.

Christophe, A., Nespor, M., Guasti, M. T., & van Ooyen, B. (2003). Prosodic structure and syntactic acquisition: the case of the head-direction parameter. *Developmental Science, 6*, 211-220.

Christophe, A., Peperkamp, S., Pallier, C., Block, E., & Mehler, J. (2004). Phonological phrase boundaries constrain lexical access: I. Adult data. *Journal of Memory and Language, 51*, 523-547.

de Pijper, J. R., & Sanderman, A. A. (1994). On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *Journal of the Acoustical Society of America, 96*, 2037-2047.

Di Cristo, A. (2000). Vers une modélisation de l'intonation du français: seconde partie. (Towards a model of French intonation: Second part). *French Language Studies, 9*, 143-179.

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.*

Fisher, C., & Tokura, H. (1996). Prosody in speech to infants: Direct and indirect acoustic cues to syntactic structure. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 343-363). Mahwah, NJ: Lawrence Erlbaum Associates.

Fougeron, C., & Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America, 101*, 3728-3740.

Frauenfelder, U. H., & Peeters, G. (1990). Lexical segmentation in TRACE: An exercise in simulation. In G. Altmann (Ed.), *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives* (pp. 50-86). Cambridge, Mass: MIT Press.

Gerken, L. (1994). Young children's representation of prosodic phonology: evidence from English-speakers' weak syllable productions. *Journal of Memory and Language, 33*, 19-38.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1*, 3-55.

Gleitman, L., & Wanner, E. (1982). The state of the state of the art. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 3-48). Cambridge UK: Cambridge University Press.

Gout, A. (2001). *Etapes précoces de l'acquisition du lexique.* Unpublished PhD thesis, Ecole des Hautes Etudes en Sciences Sociales, Paris.

Gout, A., Christophe, A., & Morgan, J. (2004). Phonological phrase boundaries constrain lexical access: II. Infant data. *Journal of Memory and Language, 51*, 548-567.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin, 75*, 424-429.

Hallé, P., Durand, C., & de Boysson-Bardies, B. (2008). Do 11-month-old French infants process articles? *Language and Speech, 51*, 45-66.

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language, 27*, 689-705.

Hayes, B., & Lahiri, A. (1991). Bengali intonational phonology. *Natural Language and Linguistic Theory, 9*, 47-96.

Hirsh-Pasek, K., Nelson, D. G. K., Jusczyk, P. W., Cassidy, K. W., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition, 26*, 269-286.

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*, 341-353.

Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press/Bradford Books.

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

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

Keating, P., Cho, T., Fougeron, C., & Hsu, C.-S. (2003). Domain-initial articulatory strengthening in four languages. In J. Local, R. Ogden & R. Temple (Eds.), *Papers in Laboratory Phonology, VI: Phonetic Interpretation* (pp. 143-161). Cambridge: Cambridge University Press.

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.

Kemler-Nelson, D. G., Hirsh-Pasek, K., Jusczyk, P. W., & Cassidy, K. W. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language, 16*, 55-68.

Kern, S. (2007). Lexicon development in French-speaking infants. *First Language, 8*, 227-250.

Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language, 40*, 153-194.

Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology-General, 134*(4), 477-500.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1-86.

McQueen, J. M., Cutler, A., Briscoe, T., & Norris, D. (1995). Models of continuous speech recognition and the contents of the vocabulary. *Language and Cognitive Processes, 10*, 309-331.

McQueen, J. M., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 621-638.

Mersad, K., Goyet, L., & Nazzi, T. (this volume). Cross-linguistic differences in early word form segmentation: a rhythmic-based account. *Journal of Portuguese Linguistics*.

Millotte, S., Frauenfelder, U., & Christophe, A. (2007). *Phrasal prosody constrains lexical access*. Paper presented at the Architectures and Mechanisms for Language Processing.

Millotte, S., René, A., Wales, R., & Christophe, A. (2008). Phonological phrase boundaries constrain on-line syntactic analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 874-885.

Millotte, S., Wales, R., & Christophe, A. (2007). Phrasal prosody disambiguates syntax. *Language and Cognitive Processes, 22*, 898-909.

Morgan, J. L. (1986). *From simple input to complex grammar*. Cambridge Mass: MIT Press.

Morgan, J. L., Shi, R., & Allopenna, P. (1996). Perceptual bases of rudimentary grammatical categories: Toward a broader conceptualization of bootstrapping. In J. L. Morgan & K. Demuth (Eds.), *Signal to Syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 263-283). Mahwah, NJ: Lawrence Erlbaum Associates.

Morgan, J. L., Singh, L., Bortfeld, H., Rathbun, K., & White, K. S. (2001). *Effects of speech rate and sentence position on infant word recognition*. Paper presented at the Boston University Conference on Language Development.

Name, C., & Silva, C. (2009). *Prosodic cues in Brazilian Portuguese second language learners*. Paper presented at the Boston University Conference on Language Development.

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

Nespor, M., & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris.

Newman, R. S., Bernstein Ratner, N., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: A retrospective analysis. *Developmental Psychology, 42*, 643-655.

Norris, D. G. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition, 52*, 189-234.

Norris, D., McQueen, J. M., & Cutler, A. (1995). Competition and segmentation in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1209-1228.

Pasdeloup, V. (1990). *Modèle de règles rythmiques du français appliqué à la synthèse de la parole (A model of rhythmical rules for French applied to speech synthesis)*. Unpublished Ph.D. Thesis, Université de Provence, Aix-en-Provence.

Salverda, A. P., Dahan, D., & McQueen, J. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition, 90*, 51--89.

Selkirk, E. O. (1984). *Phonology and Syntax: the Relation between Sound and Structure*. Cambridge Mass.: MIT Press.

Shafer, V. L., Shucard, D. W., Shucard, J. L., & 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., Cutler, A., Werker, J., & Cruickshank, M. (2006). Frequency and form as determinants of functor sensitivity in English-acquiring infants. *Journal of the Acoustical Society of America, 119*, EL61-EL66.

Shi, R., & Gauthier, B. (2005). Recognition of function words in 8-month-old French--learning infants. *Journal of the Acoustical Society of America, 117*, 2426-2427.

Shi, R., & Melançon, A. (in press). Syntactic categorization in French-learning infants. *Infancy*.

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. (2006). 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.

Shukla, M., Nespor, M., & Mehler, J. (2007). An interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology, 54*, 1-32.

Singh, L., Bortfeld, H., Rathbun, K., & Morgan, J. L. (2000). *Effects of speeded and slowed speech on infant word recognition*. Paper presented at the Boston University Conference on Language Development.

Singh, L., Morgan, J. L., & White, K. S. (2004). Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language, 51*, 173-189.

Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology, 59*(1), 67-95.

Welby, P. (2003). *The slaying of Lady Mondegreen, being a study of French tonal association and alignment and their role in speech segmentation.* Unpublished PhD thesis, Ohio State University, Columbus.

Welby, P. (2006). French intonational structure: Evidence from tonal alignment. *Journal of Phonetics, 34*(3), 343-371.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America, 91*, 1707-1717.

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*, 199-231.

Séverine Millotte
Laboratoire d'Étude d'Apprentissage et du Dévelopement, CNSR – UMR 5022, Université de Bourgogne, Dijon, France
Severine.millotte@u-bourgogne.fr

James Morgan
Department of Cognitive and Linguistic Sciences
Brown University,
Providence, Rhode Island 02912, USA
james_morgan@brown.edu

Sylvie Margules
Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS – UMR 8554/DEC – ENS/EHESS, Paris, France
sylvie.margules@ens.fr

Savita Bernal
Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS – UMR 8554/DEC – ENS/EHESS, Paris, France
savita.bernal@ens.fr

Michel Dutat
Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS – UMR 8554/DEC – ENS/EHESS, Paris, France
michel.dutat@ens.fr

Anne Christophe
Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS – UMR 8554/DEC – ENS/EHESS, Paris, France
anne.christophe@ens.fr