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
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 grinceux], a grumpy cat containing the potential competitor word ‘chagrin’ sorrow was processed more slowly than [un chat drogue] 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] [srimpai...], 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 12-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 psey] [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 ball] [conclura la raison] ‘ball’-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 ‘ball’-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) [conclusa...] or constituted one syllable of a continuing prosodic group (...balcon...). Although the presence of a phonological phrase boundary 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
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 'vipère', 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).

<table>
<thead>
<tr>
<th>PP boundary</th>
<th>No boundary</th>
<th>Difference</th>
<th>t-test</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>[con]/[per]</td>
<td>[balcon]</td>
<td>[vipère]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SE)</td>
<td>Mean (SE)</td>
<td>Mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V0</td>
<td>272 (8.4)</td>
<td>307 (11.2)</td>
<td>-35</td>
<td>2.9</td>
</tr>
<tr>
<td>V1 (a/i)</td>
<td>349 (8.3)</td>
<td>273 (8.7)</td>
<td>76</td>
<td>6.9</td>
</tr>
<tr>
<td>V2 (ê/ê)</td>
<td>357 (9.1)</td>
<td>392 (5.4)</td>
<td>-35</td>
<td>3.5</td>
</tr>
<tr>
<td>V1-V0</td>
<td>77 (10.0)</td>
<td>-34 (16.5)</td>
<td>111</td>
<td>6.2</td>
</tr>
<tr>
<td>V2-V1</td>
<td>8 (12.5)</td>
<td>119 (10.7)</td>
<td>-111</td>
<td>7.2</td>
</tr>
</tbody>
</table>

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'ai aimerait 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.
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

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
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...
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


Séverine Millotte
Laboratoire d'Etude d'Apprentissage et du Développement, CNRS – 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/HESS, Paris, France
sylvie.margules@ens.fr

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

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

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