Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Phonetics 31 (2003) 585–598

Journal of Phonetics

www.elsevier.com/locate/phonetics

Discovering words in the continuous speech stream: the role of prosody

Anne Christophe[a,*], Ariel Gout[a,b], Sharon Peperkamp[a,c], James Morgan[a,d]

[a] Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS-ENS-CNRS, Paris, France
[b] Service de Neuropédiatrie, CHU Bicêtre, Paris, France
[c] Université de Paris VIII, Paris, France
[d] Brown University, Providence, USA

Received 29 July 2002; received in revised form 27 May 2003; accepted 10 June 2003

Abstract

Finding words in sentences is made difficult by the absence of obvious acoustic markers at word boundaries, such as silent pauses. Recent experimental evidence suggests that both adults and infants are able to use prosodic boundary cues on-line to constrain lexical access. French adults performing a word detection task were slowed down by local lexical ambiguities within phonological phrases but not across a phonological phrase boundary (Christophe, Peperkamp, Pallier, Block, & Mehler, J. Mem. Language (in revision)). Thirteen-month-old American infants who were trained to turn their heads upon hearing a bisyllabic word, such as 'paper', in a variant of the conditioned head-turning paradigm, responded more often to sentences that contained the target word than to sentences containing both its syllables separated by a phonological phrase boundary (Gout, Christophe, & Morgan, J. Mem. Language (in revision)). Taken together, these results suggest that both French adults and 13-month-old American infants perceive phonological phrase boundaries as natural word boundaries, and that they do not attempt lexical access on pairs of syllables which span such a boundary. We discuss the potential generalization of these results to other languages, the universality of prosodic boundary cues as well as their use in on-line syntactic analysis and syntax acquisition.

1. Introduction

Finding words in sentences is made difficult by the frequent absence of obvious acoustic markers at word boundaries, such as silent pauses. Adult speakers of a language are often

*Corresponding author. LSCP, ENS, 46, rued'ulm, 75005 Paris, France. Tel.: +33–1–44–32–23–55; fax: +33–1–44–32–23–63.

E-mail address: anne.christophe@ens.fr (A. Christophe).

0095-4470/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0095-4470(03)00040-8

assumed to rely heavily on direct identification of known lexical items (e.g., McClelland & Elman, 1986, for the first formulation of this hypothesis). In this view, adult speakers are supposed to activate all the lexical items which are compatible with the incoming phonetic information at any moment in time; whenever a given phoneme or syllable may be attributed to more than one word, these words enter in competition (e.g., "can" in "canvass"). This process of multiple activation and competition ensures that when the system stabilizes, each phoneme or syllable is attributed to one and only one word, yielding a complete parse of the spoken sentence. Although this strategy works in practice (as shown by many computer simulations), and has even received experimental support in adults (see e.g., McQueen, Norris, & Cutler, 1994; Norris, McQueen, & Cutler, 1995), it is not initially available to infants who have to first learn the lexicon of their language.

Over the years, a number of word boundary cues have been identified that may be used by infants as well as by adult speakers. Experimental evidence has shown that infants are sensitive to all of these cues.

To review them briefly, *allophonic cues* refer to the fact that phonemes have different phonetic realizations (*allophones*) depending on their positions in words or syllables; for instance, both the /t/ and /r/ in "night rates" are different from those in "nitrates" (namely, /t/ is aspirated, released, and retroflexed in "nitrates" while it may be unaspirated and unreleased or glottalized in "night rates"; /r/ is largely devoiced in "nitrates" but voiced in "night rates"); such differences can thus be used to impute word boundaries. Two-month-old American infants have been shown to perceive these differences (Hohne & Jusczyk, 1994).

Phonotactics refer to the fact that certain phoneme sequences are impossible (or improbable) within words but possible (or probable) across a word boundary; several experiments have shown that 9-month-old infants prefer to listen to syllables or words which exemplify the most frequent phoneme sequences in their native language, thus providing evidence that they are already sensitive to their native language phonotactics (Friederici & Wessels, 1993; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993; Jusczyk, Luce, & Charles-Luce, 1994).

Infants may also use their knowledge of the *typical prosodic shape of individual words*: In English, for instance, most content words start with a strong syllable which may be followed by a weak one. Adult speakers of English have been shown to exploit this fact to hypothesize word boundaries (McQueen et al., 1994; Norris et al., 1995). Nine-month-old American infants, but not 6-month-olds, know about this property of English words: They listen longer to lists of strong–weak words (typical of English) than to lists of weak–strong words (atypical of English, Jusczyk, Cutler, & Redanz, 1993).

Finally, infants may be sensitive to the boundaries of *prosodic units*: speech is organized into a hierarchy of prosodic units. Intonational phrases often correspond to whole sentences (or propositions within a sentence), and their boundaries are marked with pitch declination, lengthening, and frequently a pause: infants as young as 4.5 months were shown to react to the disruption of intonational phrases (Hirsh-Pasek et al., 1987; Jusczyk et al., 1992). Intonational phrases are composed of one or more phonological phrases that typically contain one or two content words together with the function words that are associated with them: 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 (Kemler-Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989; Gerken, Jusczyk, & Mandel, 1994).

The studies cited above showed that infants are sensitive to each of the targeted cues, but not that they actually use them for word segmentation. Jusczyk and Aslin (1995) went one step further, and developed a novel experimental technique allowing the direct investigation of how infants access words in running speech. With this technique, Jusczyk and his collaborators began to study infants' usage of many of the above-mentioned word boundary cues, and even to compare their relative importance in cases in which several cues conflict. In brief, they observed that 7.5-month-old American infants were already able to segment bisyllabic words with the typical strong–weak stress pattern of English (as in ''kingdom''); in contrast, only 10.5-month-olds were able to correctly segment bisyllabic words with an atypical weak–strong stress pattern (as in ''guitar'', Jusczyk, Houston, & Newsome, 1999). In further experiments, Jusczyk and his colleagues showed that phonotactic and allophonic cues are also exploited by infants, and also that stress cues are weighted more heavily than either phonotactic, allophonic, or distributional cues (Jusczyk, Hohne, & Bauman, 1999; Mattys, Jusczyk, Luce, & Morgan, 1999; Johnson & Jusczyk, 2001). Jusczyk and his collaborators contributed massively to this literature. They studied many possible word boundary cues, and this is crucial because none of them is sufficient to support 100% correct segmentation on its own. However, taken together, they may allow infants to start acquiring a lexicon.

In the remainder of this paper, we focus on prosodic boundaries: just as Peter Jusczyk tested the actual use of many word boundary cues, we investigated whether prosodic boundaries are exploited on-line by both adults and infants to constrain lexical access.

2. Phonological phrase boundaries constrain on-line lexical access in French adults

As we saw in the introduction, there is already considerable evidence that major prosodic breaks, or intonational phrases, are perceived by very young infants. In addition, the available literature suggests that adults are able to exploit these prosodic boundaries on-line to constrain their syntactic interpretation of sentences (see e.g., Warren, Grabe, & Nolan, 1995; Kjelgaard & Speer, 1999). However, intonational phrases often correspond to whole propositions; as a result, few word boundaries coincide with intonational phrase boundaries. Here, we focus on the prosodic unit just below the intonational phrase in the prosodic hierarchy, namely the phonological phrase.[1] Phonological phrases typically contain one or two content words together with the some function words; they typically contain between four and seven syllables, and are characterized by preboundary lengthening (e.g., Delais-Roussarie, 1995, for French; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992, for English) and by the fact that there is one melodic contour per phonological phrase (Pasdeloup, 1990, for French; Hayes & Lahiri, 1991, for Bengali).

[1]This unit just below the intonational phrase has many different names and definitions in the literature (see, Shattuck-Hufnagel & Turk, 1996, for an excellent review); however, most authors agree in postulating at least two levels above the prosodic word, corresponding to what we call here phonological phrases and intonational phrases. Although there is disagreement about precise definitions, in all the experimental work we picked examples in which all theories clearly postulate a prosodic boundary bigger than a prosodic word but smaller than an intonational phrase.

To investigate on-line lexical access in adults listening to fluent speech, Christophe, Peperkamp, Pallier, Block, and Mehler (in revision) used a word detection task in which adults had to press a button as soon as they heard a target word in a spoken sentence. They studied the influence of a local lexical ambiguity that occurred either at a word boundary or at a phonological phrase boundary. Sentences with a local lexical ambiguity, such as '…*son chat grincheux*…' (/sɔ̃ʃagrɛ̃ʃø/) contained both syllables of a French word, e.g., '*chagrin*' (/ʃagrɛ̃/). Sentences without any ambiguity, e.g., '…*son chat drogué*…', were such that no word in French started by '*chad*…' . In this example, the competitor word ('*chagrin*') is embedded within a single phonological phrase, spanning a prosodic word boundary only (see example below, prosodic word boundary condition; square brackets mark phonological phrases). In the phonological phrase boundary condition, the competitor word spanned a phonological phrase boundary, as in '[*son grand chat*] [*grimpait aux arbres*]', which contains both syllables of the French word '*chagrin*' on each side of a phonological phrase boundary (see example, phonological phrase boundary condition).

Prosodic word boundary:

[*Le livre*] [*racontait l'histoire*] [*d'un* ***chat grin****cheux*] [*qui avait mordu*] [*un facteur*] (*chagrin*)
[*Le livre*] [*racontait l'histoire*] [*d'un* ***chat dro****gué*] [*qui dormait tout le temps*]. (**chad*…)
("The book told the story of a grumpy cat who had bitten a postman" // "sorrow")
("The book told the story of a doped cat that slept all day long").

Phonological phrase boundary:

[*D'après ma sœur*], [*le gros* ***chat***] [***grim****pait aux arbres*]. (*chagrin*)
[*D'après ma sœur*], [*le gros* ***chat***] [***dres****sait l'oreille*]. (**chad*…)
("According to my sister, the big cat climbed the trees")
("According to my sister, the big cat pricked up its ears").

Sixteen pairs of sentences were constructed for each boundary condition, and 32 native French participants took part in a word detection experiment (the target word was displayed visually before each spoken sentence was heard, e.g., 'chat' for the example above). Mean reaction times were then entered in a by-subjects and in a by-items ANOVA (F values reported as F_1 and F_2 hereafter). In the prosodic word boundary condition, Christophe and colleagues observed significantly slower response times in lexically ambiguous than in nonambiguous sentences (see Fig. 1, left-hand bars; effect size 60 ms, $F_1(1,28) = 18.1$, $p < 0.001$, $F_2(1,14) = 7.0$, $p < 0.02$). This result is congruent with the view that lexical access is slightly delayed when several words compete (in this case, '*chat*' and '*chagrin*'). In other words, acoustic/prosodic cues to word boundaries within a phonological phrase were not reliable enough to prevent the activation of candidates spanning the word boundary. Note that this does not mean that acoustic/prosodic cues were useless in that condition: simply, they were not powerful enough to allow participants to immediately close all pending lexical searches.

Now that the word detection task is established as a paradigm that is sensitive to the activation of multiple lexical candidates, the crucial experimental condition is the phonological phrase boundary condition. One of three patterns of results could obtain: First, it could be that the

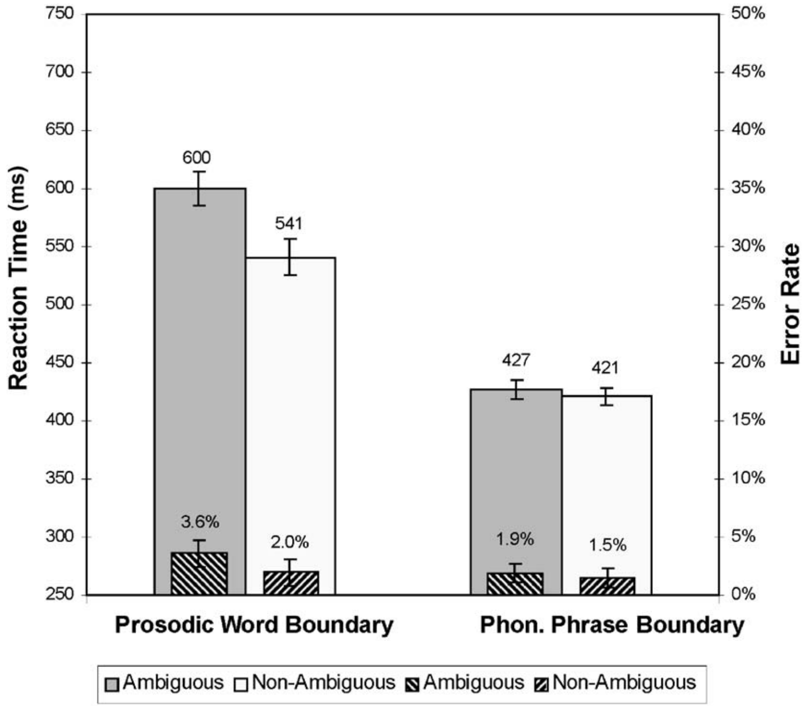

Fig. 1. Reaction times (ms) and error rates (%) in a word detection task, for sentences with or without a local lexical ambiguity, which either spanned a phonological phrase boundary or not (from Christophe et al., in revision, Exp. 2). French adults were slowed down by the local lexical ambiguity in the prosodic word boundary condition only. Error bars represent one standard error of the mean above and below the mean.

ambiguity effect would be just as strong as in the prosodic word boundary condition; in that case, the conclusion would be that phonological phrase boundaries are not exploited on-line to constrain lexical activation. Second, it could be that no ambiguity effect at all would show in that condition; the conclusion would then be that phonological phrase boundaries allow listeners to close all pending lexical searches. Third, an intermediate ambiguity effect could obtain, showing that phonological phrase boundaries are only partially exploited to constrain lexical access. As can be seen in Fig. 1 (right-hand bars), the second pattern of results was observed: in the phonological phrase boundary condition, participants were equally fast for ambiguous and nonambiguous sentences (effect size 6 ms, $F_1 < 1$, $F_2 < 1$), suggesting that a competitor word that spanned a phonological phrase boundary was never activated. Overall, there was a significant interaction between the factors boundary and ambiguity ($F_1(1, 28) = 9.8$, $p < 0.01$, $F_2(1, 29) = 4.7$, $p < 0.05$), showing that the two boundary conditions behaved differently with respect to the ambiguity effect. In addition, one may note that participants responded much faster in the phonological phrase boundary condition than in the prosodic word boundary condition (effect size 147 ms, $F_1(1, 28) = 266$, $p < 0.001$, $F_2(1, 29) = 17.8$, $p < 0.001$). In other words, the target word was detected much faster when it was immediately followed by a phonological phrase boundary. This suggests that the presence of the phonological phrase boundary facilitated lexical access to

Table 1
Mean duration of segments in the phonological phrase boundary and the word boundary condition (ms)

	Phonological phrase boundary		Word boundary		Difference	t-test		% lengthening
	Mean (ms)	S.E.	Mean (ms)	S.E.		$t(63)$	p	
S1-onset (/ʃ/)	106.4	5.8	112.2	4.5	−5.8	<1		−5.4%
S1-vowel (/a/)	111.5	5.6	79.5	2.8	32.0	5.1	$<10^{-5}$	40.2%
S1-coda (ø)	55.6	5.1	55.9	3.6	−0.3	$t(29)<1$		−0.4%
S2-onset (/gr/)	81.9	4.3	85.9	4.3	−4.0	<1		−4.8%
S2-vowel (/ɛ̃/)	65.4	2.7	85.0	3.7	−19.6	4.3	$<10^{-4}$	−30.0%
S2-coda (ø)	63.2	3.6	67.5	4.5	−4.3	$t(13)<1$		−6.8%

S1 is the first syllable involved in the ambiguity (e.g., /ʃa/ in "chat grincheux"); S2 is the second syllable (e.g. /grɛ̃/ in the same example); percent lengthening was computed as the difference in duration divided by the shortest duration and multiplied by 100.

the target word. Upon hearing the phonological phrase boundary, participants were able to close all pending lexical searches and directly identify the target word.

Given that phonological phrase boundaries had such a large effect on their behavior, adult participants must have relied on some significant acoustic/prosodic cue(s) that distinguished between sentences from both conditions (word boundary vs. phonological phrase boundary). Table 1 presents duration measurements for individual segments of both critical syllables making up the competitor word in the locally ambiguous sentences (e.g., /ʃa/ and /grɛ̃/ in "chat grincheux", S1 and S2, respectively). As expected from the literature, there was a highly significant phrase-final lengthening (e.g., Delais-Roussarie, 1995): the vowel of S1 was 40% longer when it was phonological phrase-final rather than just word-final (percent lengthening was computed as the duration difference divided by the smallest duration and multiplied by 100). The vowel of S2 was significantly longer in the prosodic word boundary condition, reflecting phrase-final lengthening as well (S2 is part of a phrase-final word in the word boundary condition, e.g., [un chat *grincheux*], while it is phrase-initial in the phonological phrase boundary condition, e.g., [*grimpait* aux arbres]). Phrase-final lengthening is thus very possibly one of the cues that adults exploited in order to compute phonological phrase boundaries. However, other cues can also influence processing, such as for instance pitch contour or the amount of coarticulation between adjacent segments (see e.g., Hardcastle, 1985; Holst & Nolan, 1995; Byrd, Kaun, Narayanan, & Saltzman, 2000). We will come back to this point in the discussion.

This experiment thus showed that phonological phrase boundaries are exploited on-line by French adults to constrain lexical access. Several studies had already shown that participants could exploit prosodic cues to segment speech into words (see e.g., Nakatani & Schaffer, 1978; Rietveld, 1980; Smith, Cutler, Butterfield, & Nimmo-Smith, 1989). What remained unclear however was when in the lexical access process this prosodic information intervened. There are two main possibilities. Either lexical segmentation is performed mainly on the basis of lexical recognition, and prosodic boundary cues are called upon to help resolve ambiguities when they arise: prosodic boundaries would be used as a last-resort strategy when lexical recognition failed. Or, the prosodic analysis of sentences might be computed in parallel with lexical activation and recognition, in which case prosodic boundaries would be one of the cues that contribute to the activation of lexical candidates. The results presented above clearly rule out the first option: since the lexical ambiguity was only local, and the next syllable permitted unambiguous lexical access, lexical recognition never failed. Nevertheless, an effect of prosodic structure was observed, with two distinct patterns of results depending on the size of the prosodic boundary. This is evidence in favor of the second option.

Christophe et al.'s experiment thus demonstrates that, at least in French, lexical access occurs within the domain of phonological phrases. In addition, we wish to stress that phonological phrase boundaries become available extremely fast: indeed, lexical candidates spanning the phonological phrase boundary were not even activated. This suggests that the computation of prosodic structure occurs very early in perceptual processing, early enough to constrain the activation of lexical candidates. As a consequence, it may also be fruitfully exploited to constrain on-line syntactic analysis. We will come back to this point in the discussion.

3. Phonological phrase boundaries constrain on-line lexical access in American infants

Our main motivation for studying nonlexical segmentation cues was that infants who are in the process of acquiring a lexicon do not have full access to the lexical segmentation strategy (which is a perfectly adequate solution for adults, as mentioned above).

How can one test whether infants also exploit phonological phrase boundaries on-line to constrain lexical access? Ideally, one would need an experimental setup allowing infants to perform a task similar to the word detection task that has been successfully used with adults. This has become possible thanks to the development of a two-session variant of the conditioned head-turning paradigm that implements word detection (in James Morgan's laboratory, at Brown University). Gout, Christophe, and Morgan (in revision) exploited this new technique and trained infants to turn their heads upon hearing a bisyllabic word, such as "paper". In a second session, they were presented with whole sentences. Some of them contained the target word "paper" (see example below, "paper"-sentence), while others contained both syllables of "paper", but separated by a phonological phrase boundary (example below, "pay][per"-sentences). In addition, some distractor sentences did not contain any syllable similar to "paper". The experimental measure was how often infants turned towards the loudspeaker, during a 2 s window that started at the beginning of the syllable /peɪ/.

"*Paper*"-*sentences*:

[*The college*] [*with the biggest* ***paper*** *forms*] [*is best*].

"*Pay*][*per*"-*sentences*:

[*The butler*] [*with the highest* ***pay***] [***per****forms the most*].

Twenty-four pairs of experimental sentences were constructed, and 24 American 13-month-olds completed both experimental sessions. Results showed that infants turned their head upon hearing "paper" in "paper"-sentences about 80% of the time (see Fig. 2). In contrast, for

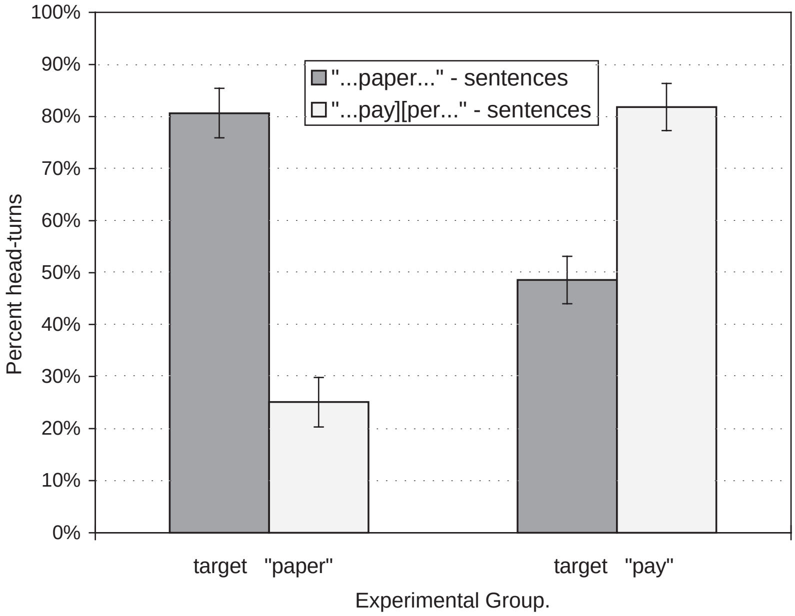

Fig. 2. Percent response to "paper"-sentences and "pay][per"-sentences for 13-month-old American infants trained to respond to either "paper" (left-hand bars) or "pay" (right-hand bars) (from Gout et al., in revision, Exp. 1). Error bars represent one standard error of the mean of the difference above and below the mean.

"pay][per"-sentences, they turned their head only about 25% of the time. This difference was highly significant (effect size 55%, $F(1,22) = 182$, $p<0.001$), suggesting that infants did not spontaneously identify the word "paper" when they heard a sentence containing both its syllables. However, this result relies on an absolute difference in response rate between two distinct sets of sentences. Infants' differential behavior might thus be attributed to some intrinsic differences between sentence types. For instance, if "paper"-sentences were more interesting for some irrelevant reason, e.g., pronounced with a higher pitch, then infants could have responded more often to these sentences, irrespective of the presence or absence of the phonological phrase boundary.

To control for this, Gout and colleagues ran another group of 13-month-olds, who were trained to turn their heads upon hearing the monosyllabic word "pay". These infants were expected to reliably respond to "pay][per"-sentences that contained the target word "pay". For "paper"-sentences the prediction was less clear, since the syllable /peɪ/ was present in the middle of a phonological phrase, and infants may have thought that it corresponded to the word "pay" (indeed, the adult results suggest that within phonological phrases, adults rely on their knowledge of words in order to identify word boundaries). However, the clear prediction was that infants trained to detect "pay" should not respond more often to "paper"-sentences than to "pay][per"-sentences: if anything, the reverse pattern should be observed. Thus, Gout and colleagues expected a significant interaction between the factors Group of infants and Sentence Type. This is indeed what they observed, since infants from the "pay" group responded to 80% of the "pay][per"-sentences, but only to 50% of the "paper"-sentences (effect size 32%, $F(1,14) = 33.1$, $p<0.001$). Crucially, the interaction between groups of infants and type of sentences was highly significant ($F(1,38) = 34$, $p<0.001$). In other words, the higher response rate to

Table 2
Mean duration of segments in "paper"-sentences and "pay][per"-sentences

	Phonological phrase boundary **pay][per**…	No-boundary **paper**	Difference		*t*-test		% lengthening
	Mean (ms)	Mean (ms)	Mean (ms)	S.E.	*t*(23)	*p*	
/p/	129	115	14	6.7	−1.9	0.06	11.4%
/eɪ/	198	112	86	10.0	8.6	$<10^{-6}$	76.4%
/p/	137	104	33	4.4	7.5	$<10^{-6}$	31.9%
/ə/	86	122	−36	5.4	6.7	$<10^{-6}$	−41.7%

Percent lengthening was computed as the difference in duration, divided by the shortest duration and multiplied by 100.

"paper"-sentences for infants in the "paper" group cannot be due to some intrinsic property of the sentences making them more interesting for infants, since infants in the "pay" group showed the opposite behavior.

This experiment thus showed that American infants, like French adults, spontaneously interpreted a phonological phrase boundary as a word boundary: they did not attempt to recognize the target word ("paper"), when its constituent syllables were separated by a phonological phrase boundary.

Measurements of the duration of the critical segments (see Table 2) revealed a highly significant phrase-final lengthening for the /eɪ/ diphthong that was 76% longer in "p*ay*]" than in "p*a*per" (note that this corresponds both to word- and phrase-final lengthening). There was also a significant lengthening (32%) of the first consonant of a word (and phrase) relative to a word-medial consonant (/p/ in "[performs" vs. "paper", this is consistent with the literature, see e.g., Quené, 1992; Fougeron & Keating, 1997; Christophe et al., 2001). Finally, the /ə/ vowel also exhibited word-final lengthening, since it was 42% longer in word-final position ("pap*er*") than in word-medial position (e.g., "p*er*forms").

One should note that, in this experiment, the same syllables occurred adjacently in the same order (100% transitional probability between "pay" and "per") and they always exhibited a strong–weak pattern. These two powerful word boundary cues (see Section 1) thus gave strong cohesiveness to the "pay" and "per" syllables. In addition, infants had heard the target word 'paper' in isolation many times during the first session (between 60 and 200 repetitions), and they were reinforced for responding to it. The target word was thus in a good position to ensure adequate lexical segmentation. Still, in these conditions, we found that infants did not access the target word when its syllables were straddling a phonological phrase boundary. It thus seems that at least in some instances, prosodic boundary cues (of the type that was studied here) are stronger than other types of word boundary cues. In other words, lexical access occurs within the domain of prosodic units that are smaller than intonational phrases.

4. Discussion

Taken together, the adult and infant results suggest that both French adults and American 13-month-olds perceive phonological phrase boundaries as natural word boundaries, and that they do not attempt lexical access on pairs of syllables which straddle such boundaries. Would these results generalize to a fully crossed design, namely would French infants behave like American infants, and American adults like French adults? Even though this should be tested experimentally, we think that the generalization is plausible. Let us consider American adults, for instance: we know that phonological phrase boundaries can be reliably perceived on-line in American English, because American infants can do so; we also know that adults may exploit prosodic boundary cues on-line to constrain lexical access, since French adults do so. As a consequence, we think it plausible that American adults would also exploit phonological phrase boundaries on-line to constrain lexical access (of course, it remains logically possible that phonological phrase boundaries would be exploited only during the initial stages of lexical acquisition in American English; adult speakers might have access to other, more efficient strategies, so that the effect of prosodic boundaries on on-line lexical access would be negligible; however, there is no evidence in favor of this possibility).

More generally, could it be the case that phonological phrase boundaries are exploited universally? Cues to phonological phrases have been measured in several unrelated languages, in particular phrase-final lengthening (e.g., Rietveld, 1980, for French; Wightman et al., 1992, for English; de Pijper & Sanderman, 1994, for Dutch; Fisher & Tokura, 1996, for Japanese; Barbosa, 2002, for Brazilian Portuguese), as well as phrase-initial consonant lengthening/strengthening (e.g., Quené, 1992, for Dutch; Keating, Cho, Fougeron, & Hsu, 2003, for Korean, Taiwanese, French, and English). Together, these two effects conspire to increase the distance between adjacent vowel onsets when they are separated by a prosodic boundary. Since vowel onsets play a major role in determining the perceived moment of occurrence of syllables (see Port, 2003; Scott, 1993), an increased interval between vowel onsets may be perceived as a rhythmic discontinuity. The temporal organization of sentences is thus potentially a powerful cue to their prosodic structure. In addition, two other cues may play a very important role, namely pitch contour and coarticulation (greater within smaller prosodic domains). Interestingly, coarticulation has been shown to influence lexical segmentation in both infants (Johnson & Jusczyk, 2001) and adults (Mattys, in revision). From these considerations, two conclusions can be drawn: First, the fact that prosodic boundary cues have been measured in a variety of unrelated languages suggests that they might well be used universally to facilitate word segmentation, although this remains to be tested experimentally. Second, it would be very interesting to study which combination of cues listeners exploit in order to compute prosodic structure on-line, in different languages. It may well be that different languages rely primarily on different types of cues (see, for instance, the discussion in Keating et al., 2003).

If listeners were indeed universally able to extract prosodic structure from speech, on-line, what would be the implications for speech processing models? As we mentioned above, prosodic boundary information would be available early enough to influence the activation of lexical items. This influence may be implemented through extra activation granted to lexical items that begin or end at a perceptible prosodic boundary, and/or inhibition for lexical items that span this boundary. The amount of extra lexical activation/inhibition would depend on how reliably a prosodic boundary would be marked. Thus, a well-marked prosodic boundary would produce total inhibition of straddling lexical candidates (as in the phonological phrase boundary condition from Christophe et al.'s adult experiment), while a less well-marked prosodic boundary may produce only partial inhibition of straddling lexical candidates (as perhaps in the prosodic word boundary condition). All in all, the availability of prosodic boundaries should decrease the number of lexical candidates activated at any moment in time, resulting in faster and more efficient lexical access. The benefits of having on-line access to prosodic boundaries appear even greater for infants. Since phonological phrases typically contain only two to three lexical items (some of which may be function words), the problem of lexical segmentation becomes much easier.

To conclude, we would like to discuss the relationship between prosodic and syntactic structure and its implications for models of comprehension. Phonological phrase boundaries depend heavily on the syntactic structure of sentences: they always coincide with syntactic phrase boundaries, even though the reverse is not true. As a result, if listeners compute a prosodic analysis of sentences (independently of lexical access) they could exploit phonological phrase boundaries to constrain their on-line syntactic analysis of sentences. In other words, the input to the syntactic analyzer would come from two sources: the lexicon that provides words, and the prosodic analyzer that provides prosodic boundaries (as well as other potentially informative aspects of prosodic structure, such as focus). For instance, in the sentences "[the great] [bear heavy loads]" and "[the great bear] [is dangerous]", 'bear' is a verb in the first sentence, but a noun in the second; there is a syntactic difference between these two sentences, that is reflected in the prosodic structure (brackets mark phonological phrases). Similar ambiguities exist in French, and recent results suggest that French adults are able to exploit phonological phrase boundaries to bias their syntactic analysis on-line (even though disambiguation was not complete, Millotte & Christophe, 2003).

The potential usefulness of prosodic structure for syntactic analysis becomes even more obvious when one considers acquisition. It has often been claimed that prosodic structure may help bootstrap syntactic acquisition (Gleitman & Wanner, 1982; Peters, 1983; Hirsh-Pasek et al., 1987; Morgan & Demuth, 1996; Nespor, Guasti, & Christophe, 1996). As mentioned above, phonological phrase boundaries coincide with syntactic boundaries, so that prosody provides partial information to syntactic structure (Morgan, 1986; Gerken et al., 1994). One should note that prosody in itself provides no cue to the labeling of constituents (into e.g., noun phrase, verb phrase, etc.). However, functional items tend to occur at the edges of phonological phrases, and they do provide labeling information. In fact, experimental evidence suggests that infants may already know some of the function words of their language by 11 months (see Shady, 1996; Shafer, Shucard, Shucard, & Gerken, 1998, for English; Höhle & Weissenborn, 2002, for German), and also that they may know about grammatical categories slightly later (Höhle & Weissenborn, 2002). Thus, a sentence such as "the little boy is running fast" might be perceived by infants as [the xxx]$_{NP}$ [is xxx]$_{VP}$, where the boundaries are given by prosody and the labeling is given by the function words, even before they know the meaning of the content words (Christophe, Guasti, Nespor, Dupoux, & van Ooyen, 1997). In fact, such a skeleton of a syntactic structure may be sufficient to facilitate the acquisition of word meanings (Gleitman, 1990). Even though these considerations are still speculative, we think that they are worth investigating.

To sum up, we reviewed two experiments suggesting that both American infants and French adults are able to compute phonological phrase boundaries on-line, early enough to influence lexical access. These results need to be generalized to other languages. If they are found to hold generally, they have implications for models of speech processing and language acquisition.

Acknowledgements

This paper is dedicated to the memory of Peter Jusczyk, who had agreed to give an invited paper at the "Temporal Integration in the Processing of Speech" conference from which this volume derives. His untimely death is a great loss for science, as well as a great personal loss for many of us.

The research reported in this paper was supported by the Human Capital and Mobility programme, the European Science Foundation (No. 255), the Action Concertée Incitative Cognitique, and the GIS Sciences de la Cognition (No. 99N35/0008), as well as by a grant from the Direction des Etudes, Recherches et Techniques to A.C. (No. 8780844), a "poste d'accueil INSERM" to A.G., a Fyssen grant to S.P., and a grant from NIH (HD 32005) to J.L.M.

References

Barbosa, P. A. (2002). Integrating prosodic and gestural timing constraints into a model of speech rhythm production. In S. Hawkins & N. Nguyen (Eds.), *Proceedings of the ISCA workshop on temporal integration in the processing of speech* (pp. 2–1). Aix-en-Provence, France, 8–10 April 2002.

Byrd, D., Kaun, A., Narayanan, S., & Saltzman, E. (2000). Phrasal signatures in articulation. In M. B. Broe, & J. B. Pierrehumbert (Eds.), *Papers in laboratory phonology V* (pp. 70–87). Cambridge, UK: Cambridge University Press.

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., 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., Peperkamp, S., Pallier, C., Block, E., & Mehler, J. (in revision). Phonological phrase boundaries constrain lexical access: I. Adult data. *Journal of Memory and Language*.

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.

Delais-Roussarie, E. (1995). *Pour une approche parallèle de la structure prosodique: Etude de l'organisation prosodique et rythmique de la phrase française.* Ph.D. thesis, Université de Toulouse-Le Mirail, Toulouse, France, unpublished.

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.

Friederici, A. D., & Wessels, J. M. I. (1993). Phonotactic knowledge of word boundaries and its use in infant speech-perception. *Perception and Psychophysics*, *54*, 287–295.

Gerken, L., Jusczyk, P. W., & Mandel, D. R. (1994). When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition*, *51*, 237–265.

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., Christophe, A., & Morgan, J. (in revision). Phonological phrase boundaries constrain lexical access: II. Infant data. *Journal of Memory and Language*.

Hardcastle, W. (1985). Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication*, *4*, 247–263.

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. (2002). The way to syntax: The prosody/lexicon/syntax interface in German infants. *Paper presented at the workshop on SLI, genes, development and cognitive disorders*, UCL, London, UK.

Hohne, E. A., & Jusczyk, P. W. (1994). Two-month-old infants' sensitivity to allophonic differences. *Perception and Psychophysics*, *56*, 613–623.

Holst, T., & Nolan, F. (1995). The influence of syntactic structure on [s] to [ʃ] articulation. In B. Connell, & A. Arvaniti (Eds.), *Papers in laboratory phonology IV: phonology and phonetic evidence* (pp. 315–333). Cambridge, UK: Cambridge University Press.

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

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

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, *64*, 675–687.

Jusczyk, P. W., Friederici, A. D., Wessels, J. M. I., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound pattern of native language words. *Journal of Memory and Language*, *32*, 402–420.

Jusczyk, P. W., Hohne, E. A., & Bauman, A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception and Psychophysics*, *61*, 1465–1476.

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

Jusczyk, P. W., Kemler-Nelson, D. G., Hirsh-Pasek, K., Kennedy, L., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, *24*, 252–293.

Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, *33*, 630–645.

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* (pp. 144–162). Cambridge, UK: Cambridge University Press.

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.

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. (in revision). Stress versus coarticulation: Towards an integrated approach to explicit speech segmentation. *Journal of Experimental Psychology: Human Perception and Performance*.

Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, *38*(4), 465–494.

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

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.

Millotte, S., & Christophe, A. (2003). Are phonological phrases exploited on-line for the syntactic analysis of spoken sentences. *Paper presented at the 16th annual CUNY conference on human sentence processing*, Cambridge, MA, USA.

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

Morgan, J. L., & Demuth, K. (1996). Signal to Syntax: an overview. In J. L. Morgan, & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 1–22). Mahwah, NJ: Lawrence Erlbaum Associates.

Nakatani, L. H., & Schaffer, J. A. (1978). Hearing words without words: Prosodic cues for word perception. *Journal of the Acoustical Society of America*, *63*, 234–245.

Nespor, M., Guasti, M. T., & Christophe, A. (1996). Selecting word order: The rhythmic activation principle. In U. Kleinhenz (Ed.), *Interfaces in phonology* (pp. 1–26). Berlin, Germany: Akademie Verlag.

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.* Ph.D. thesis, Université de Provence, Aix-en-Provence, France, unpublished.

Peters, A. (1983). *The units of language acquisition*. Cambridge, UK: Cambridge University Press.

Port, R. F. (2003). Meter and speech. *Journal of Phonetics*, *31*, doi:10.1016/j.wocn.2003.08.001.

Quené, H. (1992). Durational cues for word segmentation in Dutch. *Journal of Phonetics*, *20*, 331–350.

Rietveld, A. C. M. (1980). Word boundaries in the French language. *Language and Speech*, *23*, 289–296.

Scott, S. (1993). *Perceptual centres in speech: An acoustic analysis.* Ph.D. thesis, University College London, London, UK, unpublished.

Shady, M. (1996). *Infant's sensitivity to function morphemes.* Ph.D. thesis, State University of New York, Buffalo, unpublished.

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.

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, *25*, 193–247.

Smith, M. R., Cutler, A., Butterfield, S., & Nimmo-Smith, I. (1989). The perception of rhythm and word boundaries in noise-masked speech. *Journal of Speech, Language, and Hearing Research*, *32*, 912–920.

Warren, P., Grabe, E., & Nolan, F. (1995). Prosody, phonology and parsing in closure ambiguities. *Language and Cognitive Processes*, *10*, 457–486.

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.