INFANCY, *2*(3), 385–394

Perception of Prosodic Boundary Correlates by Newborn Infants

Anne Christophe and Jacques Mehler
Laboratoire de Sciences Cognitives et Psycholinguistique
EHESS-CNRS, Paris

Núria Sebastián-Gallés
Department of Psychology
University of Barcelona

French newborn infants were shown to perceive acoustic correlates of phonological phrase boundaries in Spanish, with stress-controlled stimuli. Access to phonological phrase boundaries may facilitate the acquisition of a lexicon, as well as some aspects of phonology and even syntax.

Acquiring a lexicon requires segmenting the continuous speech stream into words (or morphemes). In previous work, we suggested that the speech stream may be spontaneously perceived as a string of prosodic units, roughly corresponding to phonological phrases (Christophe & Dupoux, 1996; Christophe, Guasti, Nespor, Dupoux, & van Ooyen, 1997). Phonological phrases typically contain one or two content words together with some function words (e.g., [the little boy] [is running fast]; see Nespor & Vogel, 1986, for a formal definition). Finding the actual word boundaries is significantly simplified if operating on prosodic units shorter than whole utterances. First, word boundaries that coincide with prosodic boundaries are given and need not be computed through other means. Second, knowledge of these word boundaries may help infants discover some of the language-specific properties of words in their language, such as the combination of phonemes that occur at boundaries (i.e., phonotactic regularities) or typical stress patterns. This

Supplementary materials to this article are available on the World Wide Web at http://www.infancyarchives.com.

Requests for reprints should be sent to Anne Christophe, Laboratoire de Sciences Cognitives et Psycholinguistique, 54 Bd Raspail, 75270 Paris Cedex 06 France.

knowledge may then be used to discover other word boundaries not marked by prosody (see Mattys, Jusczyk, Luce, & Morgan, 1999, for phonotactics; see Cutler, 1996, and Peperkamp & Dupoux, in press, for stress). Third, procedures relying on known words or on the computation of distributional regularities between syllables or phonemes (e.g., Brent & Cartwright, 1996) have to consider a much reduced number of candidates if some boundaries are given by prosody.

In addition, prosodic structure may help bootstrap syntactic acquisition (Gleitman & Wanner, 1982; Hirsh-Pasek et al., 1987; Morgan & Demuth, 1996; Nespor, Guasti, & Christophe, 1996; Peters, 1983). Indeed, phonological phrase boundaries often coincide with boundaries of syntactic constituents. Therefore, they may provide some information as to the syntactic structure of sentences (Gerken, Jusczyk, & Mandel, 1994; Morgan, 1986).[1]

To test the plausibility of the prosodic segmentation hypothesis, Christophe, Dupoux, Bertoncini, and Mehler (1994) assessed whether French newborn infants could distinguish between French stimuli that contained a prosodic boundary or not. Bisyllabic items were extracted from spoken sentences: Half came from one long word, so that they contained no boundary whatsoever (e.g., *mati* from "mathé*mati*cien" or "cli*mati*sé"), and the other half came from the conjunction of the last syllable of one word and the first syllable of the following word (e.g., *mati* from "panora*ma ty*pique" or from "ciné*ma ti*tanesque"), so that there was a phonological phrase boundary between the two syllables (Nespor & Vogel, 1986). Christophe et al. (1994) tested 40 French 3-day-old infants with the nonnutritive sucking procedure and discovered that they could discriminate between items with and without a boundary.

These results were interpreted as showing that 3-day-old infants are sensitive to some local acoustic correlates of phonological phrases in French. Phonological phrases are typically characterized by preboundary lengthening (Delais-Roussarie, 1995; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992) and initial consonant lengthening (observed in languages such as English, Estonian, Czech, and Dutch; see, e.g., Klatt, 1974; Lehiste, 1965, 1966; Oller, 1973; Quené, 1992; Umeda, 1977). In addition, there is typically only one melodic contour per phonological phrase (Hayes & Lahiri, 1991, for Bengali; Pasdeloup, 1990, for French). Christophe et al. (1994) performed an acoustic analysis of the prosodic properties of the stimuli (i.e., duration and pitch of the individual segments) and observed both

[1]In particular, function words and morphemes tend to occur at the edges of phonological phrases. Access to phonological phrase boundaries may thus facilitate the acquisition of function words (see Shafer, Shucard, Shucard, & Gerken, 1998, for evidence that 11-month-old English babies already know some of the function words of English). Moreover, simultaneous access to both function words and phonological phrase boundaries may allow young children to start constructing syntactic analyses for sentences, even when ignoring the meaning of many content words (Christophe et al., 1997). Thus, the sentence *the little boy is running fast* may be initially perceived as $[\text{the xxx}]_{NP}$ $[\text{is xxx}]_{VP}$ (around the age of 1 year). In fact, such a skeleton of a syntactic structure may be needed to help learn the meaning of words (Gleitman, 1990).

significant initial consonant lengthening and preboundary lengthening, consistent with the existing literature. In this case, preboundary lengthening may be due in part to the fact that French has fixed word-final stress, mostly marked by duration. Thus, in *ma#ti* with a boundary, the syllable *ma* is stressed, whereas in *mati* without a boundary, both syllables are unstressed. As a consequence, newborns' discrimination of *mati* stimuli may have been enhanced by the stress difference. Although word-final stress in French may well be a reliable cue to word segmentation, such a cue does not exist for languages that do not have fixed stress on one edge of the word. In those languages, we still expect to find audible correlates of phonological phrase boundaries, independent of stress position.

To explore this issue we carried out another study using a language with flexible word stress to ensure identical stress patterns across categories. For this study, we chose Spanish, a language in which word stress falls predominantly on the penultimate syllable but can also be found in other positions.

METHOD

Stimuli

Disyllabic consonant–vowel–consonant–vowel (CVCV) items with or without a phonological phrase boundary were extracted from whole Spanish sentences (similarly to Christophe et al., 1994). In contrast to French, in both categories the first syllable was always unstressed and the second one always stressed. The items containing a phonological phrase boundary were constructed by splicing the last syllable of a polysyllabic word and the first of the following word (e.g., Manué*la tí*mida; gorí*la tí*sico); the items without a boundary were spliced out from the middle of a word containing at least four syllables (e.g., ge*latí*na; escar*latí*na). Ten words and 10 pairs of words containing the bisyllable *latí* were constructed. These words were inserted in a carrier sentence and read three times by a female native speaker of Spanish who did not know the purpose of the experiment. The recording was done in a soundproof room, and the speech signal was digitized at 16 kHz sampling frequency (8 kHz filtering, 16 bits). The speech stimuli were spliced at zero-crossings of the amplitude wave. There were 30 stimuli in each category (10 sequences repeated three times each), each stored in a separate file on the computer disk. Examples of stimuli can be found at http://www.infancyarchives.com.

Procedure

We used the high-amplitude nonnutritive sucking procedure; the whole procedure was strictly identical to the one used by Christophe et al. (1994). Stimulus presentation was made contingent on the infant's sucking. Each infant underwent one change in stimulation, either experimental (i.e., with a change in category from boundary to no boundary or vice versa) or control (i.e., without a change in category). To prevent control infants from hearing the same set of stimuli throughout, each set of stimuli was split in two subsets of 15 tokens each (i.e., 5 words repeated three times each). This procedure yielded four subsets of stimuli. Counterbalancing order of presentation yielded eight groups of participants, four experimental and four control. The switch of stimulation occurred after a predefined habituation criterion had been met: 2 consecutive min with a 25% decrement in sucking rate compared to the immediately preceding minute.

Participants were seated in a reclining chair placed in a soundproof chamber. A sterilized blind pacifier mounted on an adjustable mechanical arm and connected to a pressure transducer was used to measure the infants' sucking response. The experimenter stayed out of view behind the baby and checked that the pacifier stayed in the baby's mouth throughout the experiment. The computer recorded the pressure within the pacifier via an analog–digital card (Data Translation DT 2814) and detected the sucking responses on the basis of speed of increase and decrease in sucking and an amplitude threshold. It delivered one stimulus for each suck exceeding a fixed amplitude threshold. When the infant sucked continuously a fixed 600-msec stimulus onset asynchrony was used between two consecutive stimuli. An OROS AU22 sound board, a Rottel RA820BX3 stereo amplifier, and two Martin loudspeakers were used for the auditory output.

Participants

Participants were full-term French infants born at the Maternité Baudelocque-Port-Royal in Paris. The criteria for selection were that the infants weighed at least 2700 g at birth, had a gestational age of 38 weeks or more, had an Apgar score of 10 at 5 min after birth, and had no known hearing deficit. They were tested in their second, third, or fourth day. Forty-eight infants out of the 120 we tested served as participants. They had an average birthweight of 3,440 g (range = 2,700–4,400) and an average age of 2.6 days. Infants were excluded for crying (14), falling asleep (6), failing to habituate within 15 min (6), failing to suck (11), actively rejecting the pacifier (14), not sucking enough during the minutes around the switch (i.e., a baby who sucked less than 10 times during 2 consecutive min around the switch or who made a pause of more than 90 sec at the switch of stimulation was automatically rejected for not being stimulated enough; 16), experimenter interference (1), or technical problem (10; a leak in the pressure circuit occurred, and some infants were lost before it was fixed). After oral consent from the mother was obtained, the infants were gently awakened shortly before they were to eat, and the experiment began when they were in a quiet active state. The experiment began with a

2-min baseline phase without stimulation. Participants were randomly assigned to one of the eight conditions prior to testing.

RESULTS AND DISCUSSION

An acoustic analysis of the prosodic properties of the stimuli (i.e., duration, pitch, and energy of each individual segment) showed significant word-initial consonant lengthening (e.g., /t/ is significantly longer in the boundary than in the no boundary condition; see Table 1). This observation is consistent both with the literature and with the measurements on French. There is also a significant pitch difference: Final vowels are significantly higher pitched (e.g., /a/ is higher pitched in the boundary than in the no boundary condition; see Table 1). An analysis of the energy of segments (i.e., root mean square of the signal sampled at 16 kHz) revealed a marginal tendency for the preboundary vowel (/a/) and the postboundary consonant (/t/) to be louder than their within-word counterparts. This slightly greater amplitude may be perceived as a lengthening (Lieberman, 1960). Interestingly, we failed to observe significant preboundary lengthening in the present set of stimuli, as was the case in French (i.e., the /a/ vowel is not significantly longer in the boundary condition). However, in both French and Spanish we observed that the distance between adjacent vowel onsets is greater when a boundary occurs between them, due to lengthening of the final vowel, lengthening of the initial consonant, or both. Because perception of syllable timing is mainly linked to vowel onset (Scott, 1993), an increased distance between vowel onsets is perceived as a gap between adjacent syllables. A pilot study with adult native speakers of French showed that they were able to learn to categorize the Spanish *latí* stimuli better than chance (again, just like the French *mati* stimuli). This finding indicates that the prosodic differences we measured are perceivable by adults.

The results of the nonnutritive sucking experiment are shown in Figure 1. The habituation times, measured as number of minutes of the habituation phase, did not differ significantly for the experimental and control groups, $t(46) = 1.51$, $p > .1$. Experimental and control groups' sucking rates were not different on baseline, $t(46) < 1$, *ns,* nor over the last 3 min of the habituation phase, all three $t(46) < 1.5$, *ns*). An analysis of covariance was performed on the mean sucking rate for 4 min after the switch of stimulation, using the mean sucking rate for 2 min before the switch as a covariate. There were three between-subject factors: the experimental factor (experimental vs. control group) and two counterbalancing factors for order of presentation. There was a significant main effect of the experimental factor, $F(1,40) = 5.3$, $p < .03$; no other main effect or interaction was significant (all Fs < 1).

This experiment showed that French newborns discriminate between Spanish bisyllabic stimuli that differ in whether or not they contain a phonological phrase

TABLE 1
Acoustic Measurements

	Boundary			*No Boundary*				
Measurement	M	SD	*Range*	M	SD	*Range*	*Difference*	t *Test*, t*(58)*
Duration (msec)								
/a/	73.5	4.3	67–83	71.7	6.9	58–85	1.8	1.2
/t/	116.6	9.3	92–132	107.4	11.5	82–130	9.2	3.3**
/í/	65.3	18.5	37–114	70.0	12.8	44–92	–4.7	1.2
Whole item	294.7	20.3	258–335	285.9	21.3	239–327	8.8	1.6
Pitch (Hz)								
/a/	146	7.5	136–170	126	4.3	118–136	20	12.6**
/í/	149	11.1	133–170	153	9.7	133–174	–4	1.4
Energy (RMS)								
/a/	2224	314	1687–3395	2016	522	1376–3413	208	1.84*
/t/	195	27.7	150–254	178	36.4	131–276	16.6	1.96*
/í/	1892	279	1413–2457	1891	276	1309–2427	1	<1

Note. RMS = Root mean square of the signal sampled at 16 kHz.
*.05 < p < .10. **p < .01.

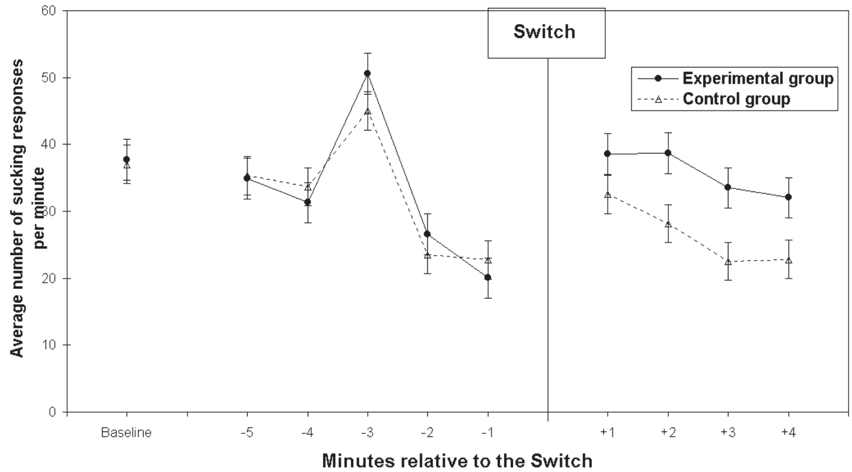

FIGURE 1 Mean sucking rates per minute for experimental and control infants. Baseline (no stimulation), last 5 min of the habituation phase, first 4 min of the test phase after the switch of stimulation. Error bars are 1 standard error above and below the mean. Infants from the experimental group (i.e., change in category) suck significantly more than infants from the control group (i.e., change of stimuli within the same category) after the change in stimulation.

boundary. This suggests that in Spanish, as in French, newborns can perceive local acoustic correlates of phonological phrase boundaries. Importantly, in the Spanish stimuli, stress always fell on the second syllable of the stimuli and therefore was not in itself a cue to distinguish between categories. Even though stress is potentially a useful cue for the perception of phonological phrase boundaries in French, it is not a necessary cue in all languages. More generally, these results suggest that correlated local cues to phonological phrase boundaries may be perceptible in a variety of languages. Are they universal or language specific? In both French and Spanish we observed a greater distance between adjacent vowels when they were on each side of a boundary as compared to when they were both within the same word. This result is also consistent with data from the literature discussed in the introduction. However, although increased distance may be a candidate for a universal boundary cue, some language-specific knowledge may be necessary to compute it. Thus, to recognize that the distance between vowels has increased, one must be able to estimate what it would have been in the absence of a boundary. This estimation may depend on language-specific properties such as presence or absence of consonant clusters, geminate consonants or long vowels, and others.

As we mentioned earlier, phonological phrase boundaries are potentially very useful in acquiring various aspects of one's mother tongue, including some aspects of phonology and syntax. In particular, lexical segmentation will be facilitated if babies have access to phonological phrase boundaries. Rather than having to parse whole utterances into words, infants are limited to finding word boundaries within strings of typically no more than seven syllables, containing one or two content words together with a few function words. Also, phonological phrase boundaries coincide with word boundaries. Recent experimental work shows that English-speaking infants start to segment words from continuous speech from the age of 7.5 months (Jusczyk & Aslin, 1995; Jusczyk, Houston, & Newsome, 1999). They also react to the disruption of phonological phrases at the age of 9 months (Gerken et al., 1994). What is needed now is direct investigation of the influence of phonological phrase boundaries on lexical access. In initial work using a word-recognition extension (Morgan, 1998) of the conditioned head-turning technique (see, e.g., Kuhl, 1983), 10-month-old infants were trained to turn their heads when hearing a bisyllabic word such as *paper* (Gout, Christophe, & Morgan, 2001). In a second session, babies were tested on whole sentences. Some of them contained the word *paper,* and babies turned their heads on hearing this word about 50% of the time. Other sentences contained both syllables of *paper* separated by a phonological phrase boundary, as in *The butler with the highest pay # performs the most*. For those sentences, babies turned their head only about 10% of the time, a highly significant difference.

To conclude, this experiment reinforces previous reports that phonological phrase boundaries may be available early in the learning process. It also helped clarify the type of information that babies may use to retrieve phonological phrase boundaries from the continuous speech stream.

ACKNOWLEDGMENTS

The work reported in this article was made possible by a grant from Direction des Recherches, Etudes et Techniques (No. 8780844) to Anne Christophe as well as by the Human Capital and Mobility programme, the Franco-Spanish Exchange Programme, and a grant from the Spanish Ministerio de Educación y Cultura (PB97-0977). We thank Jordi Pons for his help in preparing the stimuli; Caroline Floccia, Thierry Nazzi, and Brit van Ooyen for their help in testing participants; and Jim Morgan for useful comments on the revised version of this article.

REFERENCES

Brent, M. R., & Cartwright, T. A. (1996). Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition, 61,* 93–125.

Christophe, A., & Dupoux, E. (1996). Bootstrapping lexical acquisition: The role of prosodic structure. *The Linguistic Review, 13,* 383–412.

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.

Cutler, A. (1996). Prosody and the word boundary problem. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 87–99). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Delais-Roussarie, E. (1995). *Pour une approche parallèle de la structure prosodique: Etude de l'organisation prosodique et rythmique de la phrase française* [For a parallel approach to prosodic structure: A study of the rhythmical and prosodic organization of French sentences]. Unpublished doctoral thesis, Université de Toulouse-Le Mirail, Toulouse, France.

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, England: Cambridge University Press.

Gout, A., Christophe, A., & Morgan, J. (2001). *Prosodic constraints on lexical segmentation in infants*. Manuscript in preparation.

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.

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. *Journal of Cognitive Psychology, 39,* 159–207.

Klatt, D. H. (1974). The duration of [s] in English words. *Journal of Speech and Hearing Research, 17,* 51–63.

Kuhl, P. K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development, 6,* 263–285.

Lehiste, I. (1965). Juncture. *Proceedings of the Fifth International Congress of Phonetic Sciences, 1964,* 172–200.

Lehiste, I. (1966). *Consonant quantity and phonological units in Estonian* (Vol. 65, Indiana University Publications, the Uralic and Altaic Series). The Hague, Netherlands: Mouton.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America, 32,* 451–454.

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

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

Morgan, J. L. (1998, April). *Early recognition of morphophonological alternations*. Paper presented at the International Conference on Infant Studies, Atlanta, GA.

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, Inc.

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: Akademie Verlag.

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

Oller, D. K. (1973). The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America, 54,* 1235–1247.

Pasdeloup, V. (1990). *Modèle de règles rythmiques du français appliqué à la synthèse de la parole* [A model of French rhythmic rules applied to speech synthesis]. Unpublished doctoral thesis, Université d'Aix-en-Provence, Aix-Marseille, France.

Peperkamp, S., & Dupoux, E. (in press). A typological study of stress "deafness." In C. Gussenhoven & N. Warner (Eds.), *Papers in laboratory phonology* (Vol. 7). Cambridge, England: Cambridge University Press.

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

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

Scott, S. (1993). *Perceptual centres in speech and acoustic analysis*. Unpublished doctoral dissertation, University College London, England.

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.

Umeda, N. (1977). Consonant duration in American English. *Journal of the Acoustical Society of America, 61,* 846–858.

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.