# Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition

Anne Christophe, Emmanuel Dupoux, Josiane Bertoncini, and Jacques Mehler
*Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS-EHESS, Paris, France*

(Received 8 August 1992; revised 13 August 1993; accepted 22 October 1993)

Babies, like adults, hear mostly continuous speech. Unlike adults, however, they are not acquainted with the words that constitute the utterances; yet in order to construct representations for words, they have to retrieve them from the speech wave. Given the apparent lack of obvious cues to word boundaries (such as pauses between words), this is not a trivial problem. Among the several mechanisms that could be explored to solve this bootstrapping problem for lexical acquisition, a tentative but reasonable one posits the existence of some cues (other than silence) that signal word boundaries. In order to test this hypothesis, infants were used as informants in our experiments. It was hypothesized that if word boundary cues exist, and if infants are to use them in the course of language acquisition, then they should at least perceive these cues. As a consequence, infants should be able to discriminate sequences that contain a word boundary from those that do not. A number of bisyllabic stimuli were extracted either from within French words (e.g., **mati** in mathé**mati**cien), or from between words (e.g., **mati** in panora**ma ty**pique). Three-day-old infants were tested with a non-nutritive sucking paradigm, and the results of two experiments suggest that infants can discriminate between items that contain a word boundary and items that do not. It is therefore conceivable that newborns are already sensitive to cues that correlate with word boundaries. This result lends plausibility to the hypothesis that infants might use word boundary cues during lexical acquisition.

PACS numbers: 43.71.Ft, 43.71.Es, 43.71.An

## INTRODUCTION

When we listen to speech, our task is to identify the units of meaning, namely, the words (or morphemes). But, in contrast to naive intuitions, extracting words from the speech stream is not so easy: Indeed, spoken words are not flanked by silence, as blank spaces separate words in a written text. Acknowledging this fact, existing models of adult speech processing generally incorporate a lexicon-based segmentation strategy: Lexical searches are initiated at every point in the signal, activating all word candidates compatible with the acoustic information. Word recognition is achieved either through lateral inhibition of competing word candidates (as in the TRACE model, McClelland and Elman, 1986, see also Frauenfelder and Peeters, 1990) or thanks to syntactic and semantic information (as in the revised cohort model, Marslen-Wilson, 1987). Thus word boundaries may be inferred from word identity. But, even without questioning the plausibility of such a view for adult speech processing, a major problem arises when one considers language acquisition. Indeed, as Mehler *et al.* (1990) pointed out, a lexical segmentation strategy is not available to young infants. But since they nevertheless learn to speak, and in particular, they acquire a lexicon, infants presumably have to construct (acoustic, phonetic, phonemic, or other) representations of the words, and figure out the mapping between these word forms and their meanings. This implies that infants have to be able to locate words in the speech stream, even though they lack lexical knowledge.

To solve this problem, which we refer to as the bootstrapping problem for lexical acquisition, three broad categories of accounts have been proposed.

The first and most trivial proposal, states that all words to be learned are first heard in isolation. A more plausible variant of this proposal posits that a relatively large number of words are heard in isolation, enough to allow the child to bootstrap a lexical segmentation device: A fragment of speech placed between two known words would be hypothesized as being a new word.

The second type of solution states that the speech stream contains some cues (other than silence), that signal word boundaries. Moreover, infants should have the disposition to search and locate these cues. The existence of word-boundary cues has been extensively investigated by phonologists and acoustic-phoneticians, under the term "juncture phenomena." For instance, as early as 1939, Trubetzkoy described a number of potential cues that could demarcate words: He mentioned allophonic variation (e.g., aspirated unvoiced stops occur only word-initially in Tamil), lexical stress (always word-initial in Hungarian, word-final in Armenian), vowel harmony, tone phenomena, among others. Moreover, perceptual studies have shown that some of these potential cues influence adults' segmentation into words. Although these studies do not bear directly on acquisition, they add plausibility to the hypothesis that infants may rely on cues for their segmentation. Thus prosodic features such as duration, pitch, and energy were found to influence subjects' perception in several languages (see, among others, for English: Nakatani

 0001-4966/94/95(3)/1570/11/$6.00  

and Schaffer, 1978 and Cutler and Butterfield, 1990; for Dutch: Quené, 1991; for French: Rietveld, 1980). Cutler and Butterfield (1992) showed that English listeners exploit syllables' quality (strong versus weak) to hypothesize word boundaries. Finally, allophonic cues were shown to influence English subjects' parsing of pairs of words such as "gray tie" versus "great eye" (see, e.g., Nakatani and Dukes, 1977; Dutton, 1991). For the purpose of the present paper, we will group all these potential cues under the term "prelexical cues:" The generic hypothesis is that something perceivable signals word boundaries, that is represented at a prelexical level.

In contrast to the first two solutions, the third one does not postulate that word boundaries are marked in the signal (either by silence or by other prelexical cues). Instead, it assumes that word boundaries are retrieved as troughs in a transition probability function. Hayes and Clark (1970), formulated the proposal as follows: If the number of phonemes that can possibly follow a given string of phonemes is low, we probably are inside a word; if on the contrary it is high (each possible phoneme then having a low probability of occurrence), we probably are at a word boundary.

To sum up, the "words-in-isolation" hypothesis implies that the speech input received by infants is quite critical: All mothers (or caretakers) should speak to their infants uttering one word at a time. If one doesn't, her infant would fail to learn a lexicon. In contrast, the "prelexical cues" hypothesis, assumes that cues to word boundaries are normally present in speech: Mothers can, therefore, speak to their infants as they choose. However, it postulates that word boundary cues exist in every language in the world. Finally, the "transition probability" hypothesis makes no assumption at all about the form of the input.

Both the "words-in-isolation" and the "transition probability" hypotheses propose a solution to the bootstrapping problem, such that the infant becomes able to use a lexicon-based segmentation strategy. The "prelexical cues" hypothesis offers a nonlexical segmentation strategy, that could conceivably be used by adults as well as by infants.

The plausibility of these three types of accounts could be discussed at length, but because we know so little about infants' abilities, they are difficult to assess (see Christophe *et al.*, 1993). For instance, one might think that the memory load imposed by a mechanism computing transition probabilities would be too great. But for all we know, infants might very well function like tape recorders. Similarly, one might wonder how the infant decides whether a multisyllabic utterance is or is not a one-word utterance to which it should pay special attention. But for all we know, infants might make mistakes like storing two or more words as a single lexical entry, and manage to correct the error at a later stage.

Indeed, what is really needed to clarify this issue is experimental data. The "prelexical cues" hypothesis, which postulates the existence of word boundary cues in everyday speech, lends itself very nicely to experimental validation: According to this account, there has to be *something* happening around word boundaries—and whatever this something is, *infants should be able to perceive it*. Indeed, if some word boundary cues did exist, but infants were not able to perceive and represent them, they would not be able to use them to segment the speech stream. What is worse, for the case of language-specific cues, they would not be able to learn which particular cues signal words in the language they are acquiring.

Thus, for example, in French, lexical accent is said to correlate perfectly with the end of content words, which is not the case for many other languages. If infants did not even perceive whether a syllable is accented or not, how could they come to grasp that accent is a reliable marker of word boundaries? In contrast, if infants hear accent in French, then they may come to notice that every utterance they hear ends with an accented syllable, and that the accented syllable marks the end of a linguistic unit.

Thus we claim that if the "prelexical cues" hypothesis is correct, infants should be able to distinguish contexts that contain a word boundary from contexts that do not. If infants cannot discriminate such contexts, the "prelexical cues" hypothesis becomes much less plausible, whereas mechanisms that do not rely on specific prelexical cues become proportionally more attractive. However, if infants can discriminate, the hypothesis becomes plausible—yet not necessarily true; to accept the hypothesis, one would have to prove that at some point in their development, infants actually make use of such cues to segment the speech stream.

In this paper, we chose to explore newborn infants' sensitivity to word boundary cues. Of course, this sensitivity might develop in time, and it might be the case that 2-month-olds, but not newborns, perceive word-boundary cues. But, previous research has shown that newborn infants are already quite sophisticated in their processing of speech: They discriminate sentences from their native language from foreign sentences (Mehler *et al.*, 1988), they hear the difference between very short syllables that differ only in one distinctive feature (Bertoncini *et al.*, 1987), they "count" syllables (or maybe vowels, Bijeljac-Babic *et al.*, 1993). Since we assume that the ability to segment speech into word-sized units is a crucial step in language learning, it is not unreasonable to assume that newborns can perceive potential word boundary cues.

In the experiments described in this paper, we used a number of disyllabic items with the same phonemic content (/mati/). These stimuli were extracted from naturally produced spoken sentences, and differed only with respect to the presence or absence of a word boundary. That is, half of the stimuli were part of a word (e.g., "mathématicien"), while for the other half each syllable belonged to a different word (e.g., "pyjama tissé"). We investigated whether newborn infants could discriminate these two sets of stimuli. We restricted the context to two syllables under the assumption that at least some word boundary cues are local, i.e., appear near the word boundary.

## I. EXPERIMENT 1

In this experiment, we used the non-nutritive sucking procedure (Jusczyk, 1985) to assess whether three-day-old infants can discriminate bisyllabic /mati/ stimuli that either do or do not contain a word boundary. During the habituation phase, experimental infants were presented with different utterances of /mati/ chosen from a given word-boundary condition, and then switched to /mati/ drawn from the other condition; whereas control infants were then switched to a set of /mati/ physically different from that heard during the habituation phase, but belonging to the same condition. This design ensured that every single infant heard different physical tokens during the two phases of the experiment: Thus any difference of behavior between the two groups could not be due to infants noticing a change in individual tokens. Discrimination was assessed whenever the experimental group sucked more than the control group, after the switch in stimulation, but not before. In the present experiment, the statistical model of sucking rate was improved, compared to the one usually employed. A multiple regression was performed, using the mean sucking rate after the switch of stimulation as the dependent variable, and the sucking rates before the switch, as well as a dummy variable for the experimental versus control group factor as independent variables (see Appendix A for a comparison of this method to the one most often reported for the non-nutritive sucking procedure).

### A. Method

#### 1. Stimuli

The stimuli were all CVCV disyllables belonging to two different categories. The items of the first stimulus condition (BETWEEN-words) were constructed by splicing the last syllable of a polysyllabic word and the first of the following word (e.g., "pyjama tissé"); the items of the second condition (WITHIN-word) were spliced out from the middle of a word containing at least four syllables (e.g., "mathématicien"). A systematic search was conducted to find a phonemic context allowing many different occurrences of WITHINs and BETWEENs. All the disyllables that could be constructed from the consonants /b/, /d/, /g/, /t/, /p/, /m/, /n/, /r/, /l/, /s/ and the vowels /a/, /i/, /o/, /u/ were examined, using a computerized database (created from Trésor de la Langue Française, 1971).[1] The disyllable /mati/ was chosen, because it allowed us to construct 20 pairs of the type "mathématicien"/"pyjama tissé." The carrier words were inserted in a carrier sentence ("A la radio, on parle de ..., du moins je le crois."). A female native speaker of French who did not know the purpose of the experiment read all the sentences three times. She was instructed to read clearly and naturally. The recording was done in a sound-proof room, and the speech signal was digitized at a 16-kHz sampling frequency (8-kHz filtering, 16 bits). The speech stimuli were spliced using a speech editor implemented on a PDP11. We made sure that no information about the phonemes preceding /ma/ or following /ti/ could be heard. All the splicings were made at zero crossings of the amplitude wave. The stimuli that did not result in two good quality occurrences of /mati/ were excluded (the causes of exclusion were /m/ too short and heard as unnatural, or the impossibility of finding a portion of signal heard only as /i/, as in "lama tyrannique" for example). Fourteen words containing WITHINs and fourteen pairs of words containing BETWEENs were finally selected using this criterion (see Appendix B). As two occurrences of each word or pair of words were chosen, there were 28 stimuli in each category.

#### 2. Apparatus

All testing took place in a specially equipped sound-proof room at the Maternité Baudelocque (Hospital Cochin) in Paris. A sterilized blind nipple mounted on an adjustable mechanical arm and connected to a pressure transducer was used to measure the infants' sucking response. This information was fed into an electronic interface, that detected a sucking response on the basis of absolute amplitude, speed of rising, and speed of falling pressure. This interface also detected sound on the channels of the tape recorders, and sent both sets of information (sucking and sound) to a Logabax computer programmed to control the experiment. Two Tascam tape recorders, a Rottel RA820BX3 stereo-amplifier, and a Martin loudspeaker were used for the auditory output.

Stimuli were recorded continuously, in a random order, with a SOA of 600 ms. Since the longest stimulus was 410 ms long, and the shortest one 300 ms, the silent interval between two stimuli was at least 190 ms, and at most 300 ms.

The experimental program worked as follows: Each time a sucking response was recorded, the tape-recorder channel became active at the first silent interval, and stayed active for 550 ms. The channel was switched off as soon as a silent interval was found, except if the infant had sucked again during this time, in which case the channel remained active for another 550 ms. The infant received exactly one stimulus per isolated suck, with a latency between 0 and 300 ms, and heard strings of stimuli with an SOA of 600 ms when it sucked in bursts (this amounts to about as many stimuli as responses, since the average infant sucks at a rate of a little less than 2 sucks per second).

#### 3. Procedure

We used the high-amplitude non-nutritive sucking procedure, described in detail in Jusczyk (1985). This procedure allows stimulus presentation to be made contingent on infant sucking. After a baseline measurement without any stimulation, the habituation phase began. During this phase, the infant heard one stimulus from one category each time it sucked. This phase ended when there was a fall in the infant's sucking response. In this study, the predetermined fall criterion was two consecutive minutes with a 25% decrement in sucking rate compared to the preceding minute. In the second phase of the experiment, or test phase, half of the subjects switched to a different stimulus category (the experimental group), while the other half went on hearing stimuli of the same category (the control

TABLE I Layout of stimulus blocks in experimental and control groups for experiment 1

| Group | Experimental groups | | Group | Control groups | |
|---|---|---|---|---|---|
| | Phase I | Phase II | | Phase I | Phase II |
| E1 | WITHIN 1 | BETWEEN 2 | C1 | WITHIN 1 | WITHIN 2 |
| E2 | WITHIN 2 | BETWEEN 1 | C2 | WITHIN 2 | WITHIN 1 |
| E3 | BETWEEN 1 | WITHIN 2 | C3 | BETWEEN 1 | BETWEEN 2 |
| E4 | BETWEEN 2 | WITHIN 1 | C4 | BETWEEN 2 | BETWEEN 1 |

group). To prevent control infants from hearing the same set of stimuli throughout, each set of stimuli was split in two subsets of 14 tokens each. This procedure yielded four subsets of stimuli (WITHIN 1, WITHIN 2, BETWEEN 1 and BETWEEN 2). Thus we had eight groups: Four experimental groups, and four control groups. The layout of the experiment is shown in Table I.

Four tapes were recorded, one for each subset of stimuli. Once the habituation phase had begun, the whole process was automatic; the program computed the sucking rates per minute and the habituation criterion. It automatically changed channels when criterion was reached. It stored the duration between two consecutive sucks, and the moment of the switch.

#### 4. Subjects

Subjects were full-term infants born at the Maternité Baudelocque in Paris. The criteria for selection were that the infants weigh at least 2850 g, have a gestational age of 38 weeks or more, have an Apgar score of 10 at 5 min after birth and have no known hearing deficit. They were tested in their third or fourth day. Forty infants served as subjects, out of the 80 we tested. They had an average birth weight of 3390 g (range 2850 to 4090), and an average age of 3.2 days. Infants were excluded for crying (5), falling asleep (2), failing to habituate within 12 min (12), not sucking enough or too irregularly (a baby that sucked less than 10 times during two consecutive minutes around the switch was automatically rejected for not being stimulated enough—11 infants), actively rejecting the nipple during the minutes around the switch (4), or experimenter interference (6). 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. They were installed in a small reclining chair facing the loudspeaker. The mother was able to watch the experiment through a window located behind the infant. The habituation phase began when the infant was accustomed to the experimental situation, i.e., when it had sucked the nipple for 1 or 2 mins.

### B. Results

The sucking rates per minute for the experimental and control infants are shown in Fig. 1. The habituation times (number of minutes of the habituation phase) did not differ significantly for the experimental and control groups [$t(38)=1.13$, $p>0.1$]. Experimental and control groups' sucking rates were not different on baseline [$t(38)<1$], nor over the last three minutes of the habituation phase [$mn-3$ to $-1$: all three $t(38)<1$]. Following the change in stimulation, experimental infants sucked significantly more than control infants only on the third minute after the shift [$mn+1$ and $mn+2$: both $t(38)<1$, $mn+3$: $t(38)=2.17$, $p<0.05$, $mn+4$: $t(38)=1.67$, $p>0.1$].

The experimental design includes three group factors: The experimental one (experimental group versus control group), and two factors of order (order 1=beginning by WITHIN versus beginning by BETWEEN, order 2=beginning by a tape labeled 1 versus beginning by a tape labeled 2). Of these, only the experimental factor is of theoretical interest; the order factors and their interactions are only controls for potential effects of material. Thus in a first step it was verified that the order factors and their interactions did not account for a significant part of the sucking behavior after the switch of stimulation. To this end, a complete regression model including, as independent variables, all seven dummy variables that represent the three group factors, the three double interactions, and the triple interaction, was compared to a reduced model including only the relevant experimental factor. The reduced model should, in principle, be more sensitive, unless unpredicted effects of materials are evidenced. In both models, the dependent variable was the mean sucking rate for the 4 min after the switch. The sucking rates for the 2 min immediately preceding the switch were independent variables. It was found that the complete model did not account for a significantly greater part of the variance than

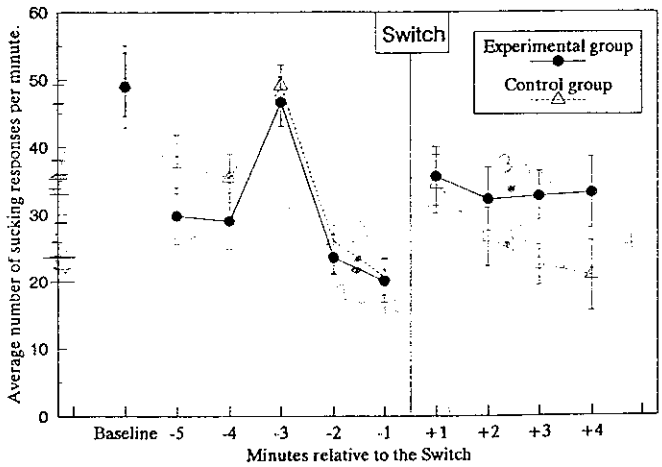


FIG. 1 Experimental 1, /mati/ stimuli: Average number of sucking responses per minute during baseline (no stimulation), 5 min before switch and 4 after. Experimental infants are switched stimulation at the moment labeled "switch," whereas control infants go on hearing stimuli from the same condition. The error bars depict one standard error of the mean above the point, and one below.

the reduced model [$F(6,30)=1.22, p>0.1$; see, e.g., Hildebrand, 1986, p. 686]. This fact justifies the use of the reduced regression model to analyze the data: The experimental factor was found to explain a significant part of the variance, with a coefficient of 8.8 [$t(36)=2.95, p<0.05$]; this means that experimental infants made, on average, 8.8 sucks per minute more than the control ones, after the switch of stimulation.[2]

### C. Discussion of experiment 1

These results show that three-days-old infants can discriminate the WITHIN-word /mati/ from the BETWEEN-word /mati/. There was no asymmetry in the response to WITHIN or BETWEEN-word /mati/, nor was there effects of materials—no effect of the two order factors.

Infants were not presented with only one item of each category, but heard a whole set of stimuli, that were extracted from different utterances. Thus the infants reacted to the differences between two sets of stimuli, and not simply to differences between two individual items. Moreover, infants could not simply have memorized all the stimuli heard during the habituation phase, and noticed that different stimuli were presented during the test phase, since control infants also heard physically different stimuli (drawn from the same category) during the habituation and the test phases. If infants did not extract a property shared by all items of one category and different between categories—if they simply reacted to novel stimuli—then control infants should have recovered as much interest in the stimuli as experimental infants, and no effect should have been observed.

One may wonder whether we observed an ability specific to young infants, or whether adult subjects also perceive a difference between stimulus categories? Indeed, adult subjects may use their lexicon to segment speech, and might have lost the infants' sensitivity to potential word boundary cues. In order to test this, the /mati/ stimuli were presented to ten adult native speakers of French in a categorization task with feedback. Nine out of ten subjects were able to categorize the stimuli better than chance ($p<0.05$) after a training phase during which they heard each stimulus once with its category label. Subjects' performance did not exceed a score of 80% correct. Furthermore, four out of five subjects performed better than chance after having received feedback over only half the stimuli. This shows that subjects can generalize what they have learned on one subset of stimuli to another subset, exactly as infants do.

In order to establish the acoustic parameters that may be responsible for carrying the distinctions, we measured the duration and energy of the two middle /mati/ segments[3] (see Table II). The duration and energy of the vowel /a/, and the duration of the /t/-closure differed significantly (all three $p<0.0001$, all other parameters measured, $p>0.01$). These values allow a correct classification of 94% of the stimuli (using a linear separation). Moreover, a multiple regression, using as a dependent variable the number of times each stimulus was classified as BETWEEN by adult subjects, and using as independent variables the three measured parameters, showed that all three parameters correlated significantly with adult behavior.

TABLE II. Duration and energy values of /mati/ stimuli: Within-word stimuli are the two middle syllable of a long word, whereas between-word stimuli are the last syllable of one word and the first of the following.

| | WITHIN | | BETWEEN | | *t* test |
|---|---|---|---|---|---|
| | mean | (s.d.) | mean | (s.d.) | *t*(54) |
| duration (ms) | | | | | |
| /a/ | 83.0 | (10.6) | 97.0 | (9.8) | 5.1** |
| /t/-closure | 55.8 | (12.6) | 74.9 | (14.0) | 5.4** |
| /t/-aspiration | 71.2 | (8.9) | 74.0 | (16.0) | 0.8 |
| energy[a] | | | | | |
| /a/ | 2216 | (379) | 2666 | (410) | 4.3** |
| /t/-closure | 117 | (37) | 98 | (29) | −2.1* |
| /t/-aspiration | 203 | (56) | 237 | (49) | 2.7* |

**$p<0.01$
*$p<0.05$
[a]Root-mean square of 16-bit sampled data

To sum up the results gathered this far, we found that newborn infants can discriminate sets of phonemically identical bisyllabic items that do or do not contain a word boundary. The design of the experiment ensured that the infants perceived some property of the stimuli that differed for the two sets of items. The adult experiment confirmed the existence of such properties, since adults were able to generalize to new stimuli. Moreover, the duration and energy measurements provided us with good candidates: The parameters measured allowed us to categorize most of the stimuli correctly, and correlated significantly with the adults' behavior.

These results, encouraging as they are, need to be replicated with better material. In experiment 1, about half of the within-word CVCV disyllables were extracted from words like "automatiquement," which in French is generally pronounced as /o/to/ma/tik/mɑ̃/. Half of our within-word CVCV stimuli were thus extracted from a CV/CVC syllabic structure, whereas for between-words stimuli, all the items belonged to a CV/CV structure. The CV/CV(C)/ stimuli were equally numerous in the WITHIN1 and WITHIN2 sets, so that any potential effect would not be revealed by the order factor order 2, in the infant study.

## II. EXPERIMENT 2

The material used in this replication was controlled for syllabic structure. To avoid any bias due to coarticulatory effects, items were matched for phonemic context of the last vowel. In addition, this replication allows generalization of the preceding results to another speaker and another phonemic context.

### A. Method

#### *1. Stimuli*

The stimuli were /mɑ̃ta/ disyllables that were extracted from words and pairs of words whose syllabic structure was always /CV/CV/. The /mɑ̃ta/ bisyllable was chosen after a computer search involving most vowels and consonants of French (i.e., /a/, /i/, /e/, /ɛ/, /o/, /u/, /y/, /ɑ̃/, /ɔ̃/, /ɛ̃/, and /p/, /t/, /k/, /b/, /d/, /g/, /v/, /z/, /ʒ/, /f/, /s/, /ʃ/, /m/, /n/, /R/, /l/) using a database created from Trésor de la Langue Française (1971). Fourteen words and pairs of words, inserted in the same carrier sentence as that used in experiment 1, along with distractors, were read three times by a female native speaker of French who did not know the aim of the experiment (she was not the same one that served in the first experiment). As there were very few mispronunciations of this disyllable, it was possible to use all three occurrences of almost each word or pair of words (words where the bisyllable was followed by /R/ had to be excluded). The stimuli were matched for the phoneme following the /a/ in the four subsets of stimuli (two subsets of WITHIN and two of BETWEEN; see Appendix C). We thus obtained three occurrences of ten words and ten pairs of words, i e , 30 stimuli per category, and 60 overall

### 2. Procedure

The apparatus and procedure were the same as in experiment 1, except that due to an improvement in the experimental program, the criterion for a fall in sucking was computed on a running window of 3 min instead of being computed only at the end of every minute, measured from the beginning of the experiment. The computer calculated the sucking rates per minute for the last 3 min at all times, and tested whether the last 2-min sucking rates were less than 66% of the antepenultimate minute. The decrement criterion was stiffened from 75% to 66%, in order to equate the mean habituation times as far as possible. Indeed, the running window procedure allowed a decrease in sucking rate to be measured at all times, so that the switch took place as soon as a 66% decrease was registered, whereas the fixed window procedure sometimes induced decreases much greater than 75% The habituation phase lasted at least 6 min. This slight modification in habituation criterion computation was thought to reduce variability between subjects, since all of them were switched as rapidly as possible, instead of being switched after a fixed number of minutes after the beginning of the experiment

### 3. Subjects

The subjects were full-term infants born at the Maternité Baudelocque. The criteria for selection were the same as for experiment 1. Infants were tested in their third or fourth day. Out of the 103 we tested, 40 infants served as subjects. They had an average birth weight of 3325 g (range 2750 to 4145) and an average age of 3.1 days Infants were excluded for crying (7), falling asleep (5), failing to habituate within 12 min (10), not sucking enough or too irregularly (a baby that sucked less than ten times during 2 consecutive minutes around the switch was automatically rejected for not being stimulated enough—11 infants), actively rejecting the nipple during the minutes around the switch (19), or experimenter interference (2). In addition to these rejection criteria that

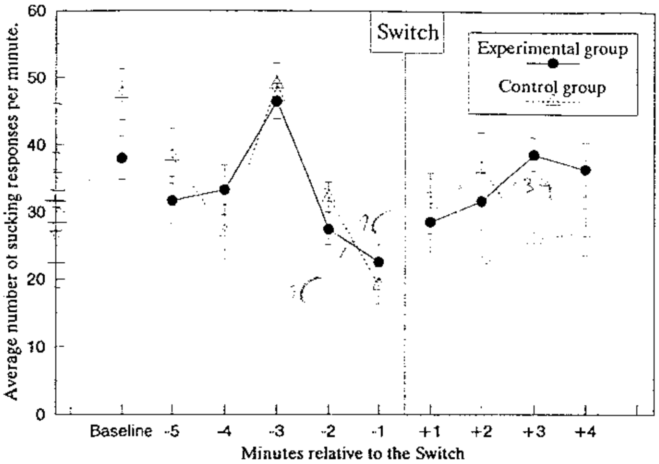


FIG 2 Experiment 2, /mɑ̃ta/ stimuli: Average number of sucking responses per minute during baseline (no stimulation), 5 min before switch and 4 after Experimental infants are switched stimulation at the moment labeled 'switch" whereas control infants go on hearing stimuli from the same condition The error bars depict one standard error of the mean above each point and one below

were already applied during experiment 1, nine more infants were rejected for sucking too much (the exact criterion was at least 1 min over 70 sucks per minute, and 2 other minutes at more than 55 sucks per minute during the habituation) Only one infant from experiment 1 would have met this criterion. This difference is probably due to a slight drift in the sensitivity of the pressure captor over time This additional criterion also accounts for the apparently higher attrition rate in this experiment (the nine infants rejected because they sucked too much met all the other criteria: 49 kept infants over 103 tested gives an attrition rate of 52%, very close to the 50% attrition rate of experiment 1)

## B. Results

The sucking patterns for the experimental and control infants are shown in Fig. 2. The habituation time (number of minutes of the habituation phase) did not differ significantly for the experimental and control groups [$t(38) < 1$]. Experimental and control groups' sucking rates were not different on baseline [$t(38) = 1.92$, $p > 0.05$], nor over the last 3 min of the habituation phase [$mn-3$ to $-1$: all three $t(38) < 1.45$, $p > 0.1$]. Following the change in stimulation, experimental infants sucked significantly more than control infants only on the third minute after the shift [$mn+1$ and $mn+2$: $t(38) < 1$, $mn+3$: $t(38) = 2.47$, $p < 0.05$, $mn+4$: $t(38) = 1.73$, $p > 0.05$].

As was the case for experiment 1, the experimental design included three group factors: The experimental one (experimental group versus control group), and two factors of order (order1 = beginning by WITHIN versus beginning by BETWEEN, order2 = beginning by a tape labeled 1 versus beginning by a tape labeled 2). Again, it was found that the complete model, with all group factors and their interactions, did not account for a significantly greater part of the variance than the reduced model, with the experimental factor only [$F(6,30) = 1.59$, $p > 0.1$]. This

**TABLE III** Duration and energy values of /mãta/ stimuli: Within-word stimuli are the two middle syllable of a long word, whereas between-word stimuli are the last syllable of one word and the first of the following

| | WITHIN | | BETWEEN | | $t$-test |
|---|---|---|---|---|---|
| | mean | (s.d.) | mean | (s.d.) | $t(54)$ |
| duration (ms) | | | | | |
| /ã/ | 71.9 | (9.7) | 85.7 | (8.9) | 5.6** |
| /t/-closure | 58.3 | (11.6) | 83.6 | (10.0) | 8.7** |
| /t/-aspiration | 19.5 | (4.6) | 17.5 | (4.1) | 1.7 |
| energy[a] | | | | | |
| /ã/ | 1176 | (387) | 1068 | (337) | −1.1 |
| /t/-closure | 176 | (83) | 101 | (46) | −4.2** |
| /t/-aspiration | 211 | (51) | 231 | (105) | 0.9 |

** $p<0.01$.
* $p<0.05$.
[a]Root-mean square of 16-bit sampled data.

fact justifies the use of the reduced regression model to analyze the data: The experimental factor was found to explain a significant part of the variance, with a coefficient of 6.3 [$t(36)=2.16$, $p<0.05$]; this means that experimental infants made on the average 6.3 sucks per minute more than the controls, after the switch of stimulation.[4]

### C. Discussion of experiment 2

These results show that three-day-old infants are able to discriminate /mãta/ disyllabic utterances that contain a word boundary from utterances that do not. The result from experiment 1, obtained with a different set of stimuli, is thus replicated. Furthermore, in experiment 2 the syllabic structure and phonemic context following the bisyllables were controlled. These two experiments allow us to conclude that, at least for French, something perceivable happens at word boundaries.

As with the /mati/ stimuli, the items were presented to adult subjects in a categorization task. Five subjects out of ten classified the two categories of stimuli better than chance ($p<0.05$), although their performance did not exceed 72% correct. Table III presents the duration and energy of the /mãta/ segments: We found that the duration of the vowel /ã/, the duration of the /t/-closure, and the energy of the /t/-closure differed significantly (all three $p<0.0001$, all other parameters $p>0.05$). These values allow a correct classification of 93% of the stimuli (using a linear separation). Overall, adult subjects did not perform quite as well in the /mãta/ as in the /mati/ categorization ($\text{chi}^2=3.8$, $df=1$, $p=0.06$). This may be attributed to greater within-condition variation in the /mãta/ stimuli characteristics, particularly changes in vowel quality and pitch pattern, that may have hampered discovery of the valid criteria. The cognitive classification task was made more difficult because salient characteristics were not useful for classification.[5]

We may wonder whether the task was more difficult with the /mãta/ rather than the /mati/ stimuli for infants, too. In order to investigate this point, the results from both experiments underwent a regression analysis, using the mean sucking rate for 4 min after switch as the dependent variable. The independent variables were the sucking rates for the 2 min before switching, the experimental factor, and a dummy variable for the interaction between the experiments (1 vs 2), and the experimental variable. The experimental variable contributed very significantly to the model, with a coefficient of 6.65 [$t(75)=3.15$, $p<0.05$], but not the interaction variable [$t(75)=1.17$, $p>0.1$]. It is thus not justifiable to consider that one contrast was more difficult for the infants than the other.

## III. GENERAL DISCUSSION

The experiments reported in this paper lend support to the idea that three-day-old infants discriminate bisyllabic stimuli extracted from naturally produced French sentences according to whether they contain a word boundary or not. Duration and energy measurements performed on the same stimuli show that the duration of the last vowel of a word, as well as the duration of the initial consonant of a word, are reliable cues for the presence of a word boundary. We also found that adult native speakers of French can learn to categorize the same stimuli.

Taken together, these findings demonstrate the existence of some perceptible differences between phonemically equivalent CVCVs spliced out from the middle of a polysyllabic word and CVCVs spliced out from between two words. This finding increases the credibility of the hypothesis that a means of segmenting words on the basis of prelexical cues exists. Further studies will be necessary before one can conclude that these potential cues for word boundaries are actually used by infants or by adults when processing continuous speech.

One may speculate about the kind of cue that was responsible for the discrimination we observed. The most salient candidate is accent. Since accent is word-final in French, between-word stimuli bear an accent on their first syllable, and within-word stimuli bear none. Hence, accent may be responsible for the observed pattern of results. If so, we may conclude that infants perceive accent in French, that they thus have the potential opportunity to discover that accent in French correlates with the end of linguistic units, and that they *might* be using accent to help segmentation, during their acquisition of French. The reader should be aware that this result was by no means predictable, for two different reasons. First, languages differ in the way they implement accent. In English, the language that was used for previous studies on accent, stress is marked by changes in vowel quality (full vowel versus schwa), as well as by differences in duration, pitch, and energy. In French, accent is marked only by differences in duration, energy, and pitch. Although some studies demonstrated that infants can discriminate English stimuli that differ only in stress pattern (see, e.g., Jusczyk and Thompson, 1978), this remained to be proven for French. Second, previous studies on accent always contrasted only two synthetic stimuli: One token with stress on the first syllable, and one with stress on the second. Thus discrimination might rely on a very shallow level of representation, e.g., comparing the length of the first syllable in the two tokens. In contrast, in the present study the infant heard many different tokens in each category, that were extracted from connected sentences. If the infants' behavior was mediated by accent, one may conjecture that this parameter could be useful in a natural setting.[6]

Lengthening of word-initial consonants is another candidate to be considered. In our stimuli, we observed a very significant lengthening of word-initial consonants (in our case always /t/), as compared to word-medial ones. Interestingly, word-initial consonant lengthening has also been reported for English (Umeda, 1977), for Dutch (Quené, 1991), for Czech (Lehiste, 1965), for Estonian (Lehiste, 1966), and for Italian and Swedish (cited in Vaissière, 1983). Such a systematic pattern in these widely different languages seems to us a potentially universal one. How infants might use word-initial consonant lengthening for segmentation has to be studied in greater depth.

The present study has important implications for models of lexical acquisition. However, the result needs to be generalized in at least two directions.

First, the study was conducted using exclusively French stimuli. The result should thus be replicated using stimuli from other languages. We mentioned above that accent is a quite plausible candidate for discrimination. But of course, accent does not systematically mark the end of content words in all languages. A first replication of the result might thus use a language where the stress pattern can be decorrelated from the presence/absence of a word boundary. The daring prediction would be that in this case (e.g., Spanish), although there is no accent difference, some other potential cue is present, that would allow discrimination.

Second, all the tokens presented in one experiment were phonemically identical, pronounced by the same speaker and at about the same rate. Any acoustic parameter could thus be directly used, without any sort of normalization. To take just one example, the duration of the *t* closure was found to be a reliable indicator of category; a very low-level measure of the amount of silence in each stimulus allows to exploit this regularity. Of course, the situation is more complex in the real world, where any segment may occur at boundaries: Here, something like the length of a segment has to be judged in relative terms, taking into account the phonemic category of this segment, the speech rate, etc. (see Klatt, 1976). Some normalization is needed. It would be possible to try and replicate the same sort of study, using many different phonemic contexts instead of only one. However, because of the habituation-dishabituation method we use, the more within-condition variability there is, the less likely infants are to identify the one feature that is shared by stimuli from the same condition, and differs between conditions. Although a positive result would certainly be very informative in this case, a negative one would be difficult to interpret.

Grosjean and Gee (1987) argued that for adults, the unit of lexical access may differ from the "written dictionary" word—i.e., everything that is flanked by two blank spaces in written text. Transposing this line of thought to infants, we can try to specify the candidates that can act as prosodic and/or linguistic units. Fortunately, the recent developments of the theory of prosodic phonology offers a whole hierarchy of prosodic units, that are the domains of phonological rules, and might well be prosodically marked in the signal. In the present study, we carefully avoided to confound a major syntactic boundary with the word boundary. We thus used noun + adjective contexts in the boundary condition. But according to the definitions of prosodic phonology theory, we have in fact a (small) phonological phrase boundary between a noun and an adjective, when the adjective follows the noun (see e.g., Nespor and Vogel, 1986; and De Jong, 1990). Thus the boundary studied in this paper is in fact a boundary between two fairly large constituents from the prosodic hierarchy. It is a question of interest to discover which unit, among the plausible candidates—e.g., phonological word, clitic group, small or maximal phonological phrase—is in fact used by infants, and whether the level perceived is universal or language specific.

## ACKNOWLEDGMENTS

This research was supported by a grant from the Direction des Recherches, Etudes et Techniques to the first author (No. 8780844) as well as by grants from the Human Frontier Scientific Program, the Human Capital and Mobility Program, and CNET. We would like to thank Anne Cutler, Stanislas Dehaene, and Christophe Pallier for their help, comments and fruitful discussion, as well as two anonymous reviewers for their thoughtful comments.

## APPENDIX A: COMPARISON OF TWO STATISTICAL MODELS FOR USE WITH THE NON-NUTRITIVE SUCKING PROCEDURE

The non-nutritive sucking procedure compares two groups of infants: Experimental infants and control infants. The two groups of infants hear stimuli belonging to the same category until a predefined habituation criterion (computed over 3 consecutive minutes) has been reached. After this point, experimental infants are switched to auditory stimuli from another category, whereas control infants go on hearing stimuli from the same category.

Discrimination is assessed whenever the experimental group sucks more than the control group after the switch of stimulation, and when this difference is not attributable to any difference before the switch of stimulation (before the switch, the two groups were treated in the same way). This test is usually made up of two steps:

Step 1: Check that the sucking rate before the switch is not different for the two groups;

Step 2: Test whether the increase in sucking rate is greater for the experimental group than for the control group.

If we call $B$ the mean sucking rate before switching, $A$ the mean sucking rate after switching, and $E$ the experimental/control factor, step 2 is a $t$-test of the effect of $E$ on $(A\text{-}B)$ (the increase in sucking rate) with the following statistical model:

$$A = B + \alpha E + k.$$

[illegible]ly writing the statistical model allows us to see one [illegible]ption that underly this test: Sucking rates after and [illegible] switch ($A$ and $B$) should be proportional, in the [illegible]ence of experimental manipulation, *with a proportional[illegible]ty coefficient of 1*. To avoid having to make this assumption, the statistical model may be modified as follows:

$$A=\beta B+\alpha E+k,$$

in this equation, the proportionality coefficient $\beta$ between $A$ and $B$ is computed, instead of being assumed to be unity. We present two reasons why this assumption is not justified, showing that the $\beta$ coefficient should be present in the statistical model. The first one is that the switch occurs at a specific moment in the sucking curve: Namely just after the habituation criterion is reached. If the switch occurred at a random moment for each baby, it could be assumed that sucking rates stay more or less constant. Since the point chosen for the switch is all but random (indeed, habituation is supposed to occur at that point, and to be reflected in sucking rate patterns), it is not justified to assume that sucking rate is identical before and after the switch in the absence of experimental manipulation (for control infants, the statistical model simplifies to $A=B+k$).

Second, even if we were to assume that the sucking rates do not show a systematic difference before and after the switch, it may be demonstrated that the slope of the regression line between $A$ and $B$ is *less than 1*. Indeed, in this case $A$ and $B$ may be considered to be two measurements of a constant "theoretical" sucking rate, $S$, taken at two different moments with some amount of error $\varepsilon$. That is

$$B=S+\varepsilon_B,$$

$$A=S+\varepsilon_A,$$

where $S$ is a random variable representing theoretical sucking rate, and $\varepsilon_B$, $\varepsilon_A$ are two error terms that may be hypothesized to follow the same distribution with mean 0 and variance $\sigma_\varepsilon^2$. In these conditions the slope of the regression line $\beta$ in

$$A=\beta B+k$$

is

$$\beta=\frac{\Sigma\, b_i a_i}{\Sigma\, b_i^2}=\frac{\Sigma(s_i+\varepsilon_{bi})(s_i+\varepsilon_{ai})}{\Sigma(s_i+\varepsilon_{bi})^2},$$

$$\beta=\frac{\Sigma\, s_i^2+\Sigma s_i(\varepsilon_{bi}+\varepsilon_{ai})+\Sigma\varepsilon_{bi}\varepsilon_{ai}}{\Sigma\, s_i^2+2\Sigma\, s_i\varepsilon_{bi}+\Sigma\varepsilon_{bi}^2}$$

Under the assumption that the error terms $\varepsilon_B$ and $\varepsilon_A$ are independent of one another as well as of $S$, all double products sums may be assumed to be vanishingly small compared to the sums of squares. Thus an approximation of the slope, $\beta$, is

$$\beta\approx\frac{\Sigma\, s_i^2}{\Sigma\, s_i^2+\Sigma\varepsilon_{bi}^2}=\frac{\Sigma\, s_i^2}{\Sigma\, s_i^2+\sigma_\varepsilon^2}.$$

According to this, $\beta$ is normally *less than 1*, and would approach 1 only if the variance of the error term, $\sigma_\varepsilon^2$, was close to zero—a rather unlikely assumption (indeed, an estimation of $\sigma_\varepsilon^2$ is the variance of the sucking rates per minute for a given baby over time, and this is typically fairly large).

An intuitive illustration of the implications of this last point is that if the error in measuring $B$ was very large, there would be little information added by $B$ in the model. Using $A$-$B$ as the dependent variable could lower the sensitivity of the test, since $B$ would only add noise. In this case, the correlation coefficient between $A$ and $B$ would be small and $B$ would have little weight in the more complex model.

To sum up, we argued that the statistical model of sucking rate that has been employed in most infant studies with non-nutritive sucking makes an unwarranted assumption, that the proportionality coefficient between sucking rates after and before switch is equal to 1. We offer instead a slightly more complex statistical model, of the following form:

$$A=\beta B+\alpha E+k.$$

This more accurate model should yield more sensitive statistics. In practice the test is a multiple regression analysis with $A$ as dependent variable, $B$ and $E$ as independent variables (the ANOVA or $t$-test usually realized were logically equivalent to a multiple regression, with ($A$-$B$) as the dependent variable, and $E$ as the independent variable).

## APPENDIX B: EXPERIMENTAL MATERIAL: /mati/ STIMULI

| Within-word stimuli: | Between-words stimuli: |
|---|---|
| WITHIN 1 | BETWEEN 1 |
| affirmativement | puma timoré |
| aromatisé | amas tissulaire |
| systématiquement | panama tisane |
| flegmatiquement | tréma titubant |
| cinématiquement | anonymat titularisé |
| grammaticalement | sigma typographié |
| mathématicien | format ticket |
| WITHIN 2 | BETWEEN 2 |
| approximativement | lama timide |
| climatisé | pyjama tissé |
| schématiquement | plasma tisonné |
| dogmatiquement | primat titré |
| diplomatiquement | cinéma titanesque |
| grammatical | coma typique |
| mathématicien | schéma tigré |

Two repetitions of each word, i.e., 14 stimuli per subset, and 56 overall

## APPENDIX C: EXPERIMENTAL MATERIAL: /mɑ̃ta/ STIMULI

| Within-word stimuli: | Between-words stimuli: |
|---|---|
| WITHIN 1 | BETWEEN 1 |
| augmentation | calmant tassé |
| segmentation | condiment tachant |
| sentimentalité | garnement taloché |
| commentateur | jument tatouée |
| momentanément | élément tannant |
| WITHIN 2 | BETWEEN 2 |
| réglementation | serment tacite |
| sédimentation | ciment taché |
| instrumentaliste | déguisement talentueux |
| fomentateur | jugement tatillon |
| momentanée | piment tamisé |

Three repetitions of each word, i.e., 15 stimuli per subset, and 60 overall. The items are matched for the phoneme following the /a/.

[1] This particular version of the database comprised only an alphabetic list of words; for this reason, phonemes that are written in many different ways, such as /k/ were not included in this search.

[2] The usually reported statistics were conducted on the same data, for comparison's sake. The increase in sucking rate (mean of the first four minutes after switch, minus mean of the last two minutes before switch) was submitted to an analysis of variance, with the three between-subjects factors already mentioned: experimental, order 1, and order 2. The two order factors showed no main effect: order 1: $[F(1,32)=3.9, p>0.05]$, order 2: $[F(1,32)=1.1, p>0.1]$. There was a significant main effect of the experimental factor $[F(1,32)=9.9, p<0.05]$. None of the interactions was significant [all $F(1,32)<2$]

[3] Values for the extreme segments, /m/ and /i/ are not very representative, since the phonetic context of these two segments varied across stimuli, and the splicing was done so as to avoid hearing the neighboring phoneme, thus deleting part of the segments themselves.

[4] The usually reported statistics were conducted on the same data, for comparison's sake. The increase in sucking rate (mean of the first 4 min after switch, minus mean of the last two minutes before switch) was submitted to an Analysis of Variance, with the three between-subjects factors already mentioned: Experimental, order 1 and order 2. The factors of order showed no main effect: order 1 $[F(1,32)=3.2, p>0.05]$, and order 2 $[F(1,32)=1.3, p>0.1]$. There was no significant main effect of experimental versus control $[F(1,32)=1.9, p>0.1]$. None of the interactions was significant [all $F(1,32)<1$]. The usual statistics thus give no significant results for experiment 2; this may be taken as a cue that the more complex statistical model that is advocated in Appendix A is indeed more sensitive.

[5] Two factors induced a greater spurious variability in the /mɑ̃ta/ stimuli, compared to the /mati/ ones. First, the /mati/ speaker read in a somewhat monotonous way, yielding pitch patterns that were very uniform across occurrences. The /mɑ̃ta/ speaker, on the contrary, put more intonation in her sentences, that induced more variability. Second, contrary to the vowel /i/, the vowel /a/ is prone to changes in vowel quality due to coarticulation with the following phoneme (it is heard as /ε/ in certain contexts)

[6] An anonymous reviewer suggested that "accent-based discrimination by newborns does not justify any conclusions regarding word segmentation (and bootstrapping). It is merely accidental that the two phenomena are related in the researchers' native tongue, but the subjects tested are unaware of this relation." We are sensitive to this reviewer's concern. However, we do not share his conclusions. We would like to claim that this divergence is due more to form than to content. Indeed, we do not claim that infants perceive word boundaries *as such*. Rather we show that infants respond to some French *word-boundary correlates*. Our results do not allow us to say whether the infants' behavior is determined by accent rather than by a complex package of cues that happen to signal word boundaries in French. Moreover, we assume that infants can discriminate a number of such acoustic correlates of word-boundaries implemented in other languages even though we acknowledge that this remains to be shown. Our working hypothesis is that only a limited number of word-boundary markers are implemented by natural languages and that the task of the infant is to select among these the ones that are relevant to the surrounding language (something they could never succeed in doing unless they at least perceive these prelexical acoustic cues)


Bertoncini, J., Bijeljac-Babic, R., Blumstein, S., and Mehler, J. (1987) "Discrimination of very short CV syllables by neonates," J. Acoust. Soc. Am. **82**, 31–37

Bijeljac-Babic, R., Bertoncini, J., and Mehler, J. (1993). "How do four-day-old infants categorize multisyllabic utterances?" Dev. Psychol. **29**, 711–721.

Christophe, A., Dupoux, E., and Mehler, J. (1993). "How do infants extract words from the speech stream? A discussion of the bootstrapping problem for lexical acquisition," in *Proceedings of the 24th Annual Child Language Research Forum*, edited by E. V. Clark (CSLI, Stanford), pp. 209–224.

Cutler, A. and Butterfield, S. (1990) "Syllabic lengthening as a word boundary cue," in *Proceedings of the 3rd Australian International Conference on Speech Science and Technology*, pp. 324–328

Cutler, A. and Butterfield, S. (1992) "Rhythmic cues to speech segmentation: evidence from juncture misperception," J. Memory Lang. **31**, 218–236.

De Jong, D. (1990). "The syntax-phonology interface and French liaison," Linguistics **28**, 57–88

Dutton, D. L. (1991) "The role of allophonic variation in the perception of word boundaries," Ph.D. thesis, State University of New York at Buffalo, Buffalo, NY

Frauenfelder, U. H., and Peeters, G. (1990). "Lexical segmentation in TRACE: an exercise in simulation," in *Cognitive Models of Speech Processing*, edited by G. Altmann (MIT, Cambridge, MA), pp. 50–86

Grosjean, F., and Gee, J. P. (1987). "Prosodic structure and spoken word recognition," Cognition **25**, 135–155

Hayes, J. R., and Clark, H. H. (1970) "Experiments on the segmentation of an artificial speech analogue," in *Cognition and the Development of Language* edited by J. R. Hayes (Wiley, New York), pp. 221–234.

Hildebrand, D. K. (1986) *Statistical Thinking for Behavioral Scientists* (PWS, Boston).

Jusczyk, P. W. (1985). "The high-amplitude procedure as a methodological tool in speech perception research," in *Infant Methodology*, edited by G. Gottlieb and N. A. Krasnegor (Ablex, Norwood, NJ), pp. 195–222.

Jusczyk, P. W., and Thompson, E. J. (1978). "Perception of a phonetic contrast in multisyllabic utterances by two-month-old infants," Percept. Psychophys. **23**, 105–109.

Klatt, D. H. (1976). "Linguistic uses of segmental duration in English: Acoustic and perceptual evidence," J. Acoust. Soc. Am. **59**, 1208–1221.

Lehiste, I. (1965). "Juncture," in *Proceedings of the Fifth International Congress of Phonetic Sciences*, Münster (Karger, Basel), pp. 172–200.

Lehiste, I. (1966). *Consonant Quantity and Phonological Units in Estonian* (Mouton, The Hague).

Marslen-Wilson, W. D. (1987). "Functional parallelism in spoken word-recognition," Cognition **25**, 71–102.

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

Mehler J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., and Amiel-Tison, C. (1988). "A precursor of language acquisition in young infants," Cognition **29**, 143–178

Mehler, J., Dupoux, E., and Segui, J. (1990). "Constraining models of lexical access: the onset of word recognition," in *Cognitive Models of Speech Processing*, edited by G. Altmann, (MIT, Cambridge, MA), pp. 236–262

Nakatani, L. H., and Dukes, K. D. (1977). "Locus of segmental cues for word juncture," J. Acoust. Soc. Am. **62**, 714–719.

Nakatani, L. H., and Schaffer, J. A. (1978). "Hearing 'words' without words: prosodic cues for word perception," J. Acoust. Soc. Am. **63**, 234–245.

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

Oller, D. K (1973). "The effect of position in utterance on speech segment duration in English," J Acoust. Soc. Am **54**, 1235–1247

Quené, H. (1991) "Word segmentation in meaningful and nonsense speech," in *Actes du XIIème Congrès International des Sciences Phonétiques*, Août 1991 (Université de Provence, Aix-en-Provence, France), pp. 82–85.

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

Trésor de la Langue Française (French language frequency counts.) (1971). Nancy, France: Centre National de la Recherche Scientifique.

Trubetzkoy, N S. (1939). *Principles of Phonology*, translated by C. A M Baltaxe (University of California, Berkeley and Los Angeles, 1969); originally published in German, Grundzüge der Phonologie, Travaux du Cercle Linguistique de Prague, 7.

Umeda, N (1977) "Consonant duration in American English," J Acoust Soc. Am. **61**, 846–858.

Vaissière, J. (1983) "Language-independent prosodic features," in *Prosody: Models and Measurements*, edited by A. Cutler and D Ladd (Springer-Verlag, Berlin), pp 53–66.