To appear in: Annual Review of Language Acquisition, Vol. 2, 2002

Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues

Franck Ramus
Laboratoire de Sciences Cognitives et Psycholinguistique (EHESS/CNRS), Paris, France

Speech rhythm has long been claimed to be a useful bootstrapping cue in the very first steps of language acquisition. Previous studies have suggested that newborn infants do categorize varieties of speech rhythm, as demonstrated by their ability to discriminate between certain languages. However, the existing evidence is not unequivocal: in previous studies, stimuli discriminated by newborns always contained additional speech cues on top of rhythm. Here, we conducted a series of experiments assessing discrimination between Dutch and Japanese by newborn infants, using a speech resynthesis technique to progressively degrade non-rhythmical properties of the sentences. When the stimuli are resynthesized using identical phonemes and artificial intonation contours for the two languages, thereby preserving only their rhythmic and broad phonotactic structure, newborns still seem to be able to discriminate between the two languages, but the effect is weaker than when intonation is present. This leaves open the possibility that the temporal correlation between intonational and rhythmic cues might actually facilitate the processing of speech rhythm.

Key-words: newborn speech perception language discrimination rhythm intonation prosody bootstrapping.

Language acquisition is a field notorious for its bootstrapping problems: in essence, it seems impossible to explain how each component of language is learnt without appealing to previous knowledge of other components. How does the child learn syntax? By relying on his/her knowledge of words, their meaning, and the meaning of whole sentences, as revealed by observation (this is *semantic bootstrapping*; Pinker, 1984). But how does the child learn the meaning of words? You have to assume some notions of syntax (this is *syntactic bootstrapping*; Gleitman, 1990).

These apparent paradoxes have raised interest in the study of the raw input available to the child, i.e., the speech signal, and of how much information can be extracted thereof. In this line, Gleitman and Wanner (1982) had already long ago suggested that prosody (rhythm, intonation) might play an important role in the acquisition of syntax (this was *prosodic bootstrapping*). Prosody has also been shown to be an important cue to the segmentation of continuous speech into words (see Jusczyk, 1997 for a review). More generally, there is of course no reason to restrict the range of potential cues to prosody. *Phonological bootstrapping*, although a misnomer, now sums up the idea that a direct surface analysis of all sorts of acoustic and phonetic cues should be of great help to the first steps of phonological, lexical and syntactic acquisition (Morgan & Demuth, 1996; Christophe, Guasti, Nespor, Dupoux, & Ooyen, 1997). The present paper is dedicated to the study of one such cue, speech rhythm.

Rhythm may be viewed as a parameter that shows variation across the languages of the world. Three types of rhythm have been identified, leading to a classification of languages into three classes (Pike, 1945; Abercrombie, 1967; Ladefoged, 1975): stress-timed languages, including most Germanic languages as well as Russian, Arabic or Thai, syllable-timed languages, including most Romance languages as well as Turkish or Yoruba, and mora-timed languages, including Japanese. Historically, stress- and syllable-timing referred to the idea that stresses or syllables would be regularly paced in the corresponding languages, but this intuition was never supported by firm empirical evidence (e.g., see Roach, 1982; Dauer, 1983). However, more recent studies have shown that this rhythm typology can be grounded in languages' phonological properties (Dauer, 1987) and in acoustic-phonetic measurements (Ramus, Nespor, & Mehler, 1999; Low, Grabe, & Nolan, 2000).

In order to know whether rhythm plays a part in language acquisition, one needs to ask: (a) what other component of language rhythm might help learn, (b) whether infants are able to represent rhythmic differences, and (c) whether infants do actually use rhythm to learn that component. The first question has already been the subject of many specu-

---

I thank Jacques Mehler, Marc Hauser, Anne Christophe, Christophe Pallier, Emmanuel Dupoux and Ghislaine Dehaene-Lambertz for useful discussions, Sylvie Margules and Renate Zangl for help testing the subjects, Jacques Mehler, Anne Christophe, John Morton and Sarah Griffiths for comments on a previous version of this paper, Xavier Jeannin and Michel Dutat for technical assistance, and the Délégation Générale pour l'Armement for financial support.

Experiments 1 and 2 have already been partially reported in Ramus, Hauser, Miller, Morris, and Mehler (2000) within a different context, and are described here in greater detail with permission from the American Association for the Advancement of Science.

Correspondence to: Franck Ramus, LSCP, 54 boulevard Raspail, 75006 Paris, France. E-mail: ramus@lscp.ehess.fr.

lations. Generally speaking, if the infant could determine very early on that her native language belongs to one of three rhythm classes, this would divide roughly by three the size of the search space for the correct grammar. But more concrete hypotheses have been proposed, relying on evidence that other interesting aspects of language are congruent with the rhythm classes.

It has been argued that adult listeners perform a pre-lexical segmentation of the speech stream, in order to obtain representation units that are best suited to their native language: stress units (feet?) for English, syllables for French, and morae for Japanese (Cutler & Mehler, 1993). If this is true, then one may want to know how the child learns which representation unit is best suited to her language. The correspondence between pre-lexical unit and rhythm (verified on those three languages so far) suggests that sensitivity to speech rhythm might be the answer.

Another case of such a correspondence has arisen from studies of adaptation to fast speech. Mehler et al. (1993) have shown that, when exposed to time-compressed sentences, listeners adapt quite quickly and soon recover asymptotic comprehension. More remarkable is the finding that exposure to compressed speech in a foreign, unknown language, improves one's comprehension of compressed speech in the native language. Indeed, Spanish listeners benefit from exposure to compressed Catalan, and English listeners benefit from exposure to compressed Dutch (Pallier, Sebastian-Galles, Dupoux, Christophe, & Mehler, 1998). However, cross-linguistic adaptation does not work across the board, as shown by the failure of French listeners to benefit from exposure to compressed English. Further cross-linguistic work suggests that transfer of adaptation to fast speech operates between languages that belong to the same rhythm class, but not across different rhythm classes (Sebastián-Gallés, Dupoux, & Costa, 2000). This prompted these authors to hypothesize that speech rate normalization involves pre-lexical processes that differ across languages, in a way that is congruent with rhythm classes. Therefore, sensitivity to speech rhythm might enable the infant to select the correct speech rate normalization strategy very early on, which would be of enormous help since acoustic variability is one of the main difficulties in the way of the acquisition of stable lexical representations.

A third parameter has been shown to vary with rhythm: syllable structure. Indeed, stress-timed languages have complex syllables, syllable-timed languages less so, and mora-timed languages a very simple syllable structure. Many phonological phenomena suggest that syllable structure is not just a fact about how words are in a particular language, but a piece of abstract phonological knowledge that plays an active role in speech production and perception (Blevins, 1995; Prince & Smolensky, 1993). To cite just one example, Japanese listeners tend to "hallucinate" vowels in the middle of consonant clusters that are incompatible with Japanese syllable structure (Dupoux, Kakehi, Hirose, Pallier, & Mehler, 1999; Dupoux, Pallier, Kakehi, & Mehler, 2001). Contrary to intuition, it is not obvious for the child to learn the correct syllabic grammar. Indeed, compiling all the syllable types would require prior syllabification of the speech stream, which is itself dependent on the syllabic grammar. And compiling the codas and onsets from utterances' boundaries would not be enough, since legal codas and onsets are often different at word boundaries and word-internally. We have therefore proposed that rhythm might provide part of the missing information (Ramus et al., 1999).

A link has also been proposed between rhythm and word learning. Indeed, a large literature has documented the use of prosodic cues for lexical access. For instance, the fact that most English words are stressed on their first syllable leads listeners to expect a word boundary before stresses (Cutler, 1996). Convincing evidence has been provided that this very strategy may underlie the first steps of lexical acquisition in English (Jusczyk, Houston, & Newsome, 1999). But other strategies are needed in languages that have another or no predominant stress pattern. The infant therefore first needs to learn the correct word segmentation strategy. Again, it has been proposed that rhythm might be the cue to the correct strategy (Jusczyk, 1997). In this case, however, it remains particularly unclear how good the mapping between rhythm class and segmentation strategy is: French and Spanish are both syllable-timed, but the former has systematic word-final stress, and the latter a predominant penultimate stress pattern.

The various hypotheses described above are not mutually exclusive. On the contrary, they seem to be converging on the idea that rhythm introduces the infant to pre-lexical phonology, that is, to the representation of the speech signal that is best suited to her native language. Rhythm may thus be a key element of phonological bootstrapping: it won't necessarily get the child words or syntactic constituents or categories, but it may provide a preliminary toolbox for further, more focused analyses. In essence, it may act as a guide to the best bootstrapping strategy. Even though the present hypotheses are only tentative at this stage, both their diversity and convergence suggest that it is relevant to move on to the next question, that is, whether infants are sensitive to speech rhythm, which is the focus of the present study.

Speech rhythm perception by infants: Available evidence

Some researchers have directly studied rhythm perception by infants. Demany, McKenzie, and Vurpillot (1977) showed that 2-3 month-old infants are able to discriminate sequences of tones differing in temporal organization. Fowler, Smith, and Tassinary (1986) moreover showed that 3-4 month-olds were able to discriminate sequences of syllables whose *Perceptual-centers*[1] (Morton, Marcus, & Frankish, 1976) were isochronous or not. However, it is not clear at all whether the notion of rhythm investigated in those studies has anything to do with that of linguistic rhythm, as defined above.

Another set of studies has suggested that newborns do classify languages according to their rhythm, as revealed

---

[1] A perceptual-center is the perceived moment of occurrence of a syllable.

by their capacity to discriminate between different languages (Mehler et al., 1988; Nazzi, Bertoncini, & Mehler, 1998; Ramus et al., 2000). The evidence proposed is twofold. First, one may look at the pattern of successes and failures accumulated across different experiments using different language pairs: newborns have been shown to discriminate between stress-timed and syllable-timed languages (French-Russian and English-Italian: Mehler et al., 1988; English-Spanish: Moon, Cooper, & Fifer, 1993) and between stress-timed and mora-timed languages (English-Japanese: Nazzi et al., 1998; Dutch-Japanese: Ramus et al., 2000). However, the only attempt to assess within-class discrimination has yielded a negative result (English-Dutch: Nazzi et al., 1998)[2]. Moreover, newborns also discriminated between a set of English and Dutch (stress-timed) sentences and a set of Spanish and Italian (syllable-timed) sentences, but failed when the two sets (English and Italian versus Dutch and Spanish) did not reflect two types of rhythm (Nazzi et al., 1998). Thus, they seem to be able to discriminate between sets of languages, if and only if these sets are congruent with different rhythm classes. As impressive as this result may be, the small number of languages studied does not guarantee that this pattern would hold for other unrelated languages. Indeed, considering the great variety of cues present in speech, other properties than rhythm may have allowed discrimination, and may therefore be considered as confounding factors. This concern has led these researchers to look for a second line of evidence, by reducing the speech cues that were available for discrimination. Thus, Mehler et al. (1988) successfully replicated their experiments after low-pass filtering their stimuli at 400 Hz. This process, which eliminates the higher frequencies of speech, and therefore most of the phonetic information, is thought to preserve only its prosodic properties (rhythm and intonation). Similarly, the experiments by Nazzi et al. (1998) used filtered speech exclusively, and Ramus et al. (2000) used sentences that were resynthesized in such a way as to preserve only prosodic cues (see below). Thus, there is converging evidence that prosody is all newborns need to discriminate languages.

Nevertheless, prosody does not reduce to rhythm. It remains possible that its other major component, intonation, plays a role in the observed discriminations. Although we do not know of typological studies of intonation that would allow us to make specific predictions for all the pairs of languages considered, it is, for instance, predictable that English and Japanese should be discriminable on the basis of their intonation. Indeed, English is a Head-Complement language, whereas Japanese is Complement-Head (for example, relative phrases come after the corresponding verb in English, but before it in Japanese), and this syntactic parameter is said to have a prosodic correlate, *prominence*, which is signaled both in terms of rhythm and intonation (Nespor, Guasti, & Christophe, 1996). Moreover, there is empirical evidence that some languages are discriminable purely by their intonation, including English and Japanese (Ramus & Mehler, 1999), English and French (Maidment, 1976, 1983) and English and Dutch (Pijper, 1983). In order to assess whether newborns actually perceive linguistic rhythm, it is therefore necessary to get rid of the intonation confound, that is to go beyond speech filtering and remove intonation from the stimuli[3].

Ramus and Mehler (1999) have adapted a technique, speech resynthesis, to selectively degrade the different components of speech, including rhythm and intonation. This technique has notably been used to resynthesize different versions of English and Japanese sentences, and assess which components of speech were sufficient for discrimination of the two languages. The different versions included (a) broad phonotactics + prosody, (b) prosody, (c) rhythm only, and (d) intonation only. Results showed that pure rhythm was sufficient for French adult subjects to discriminate between the two languages. Pure intonation was also sufficient, but the task seemed to be much more difficult. In the present series of experiments, we wish to apply the same rationale to the study of language discrimination by newborns, i.e., progressively eliminate the speech cues available for discrimination, and finally assess whether linguistic rhythm is, as hypothesized, the critical cue. Experiment 1 will assess the discrimination of Dutch and Japanese with all cues present, Exp. 2 the discrimination of those same sentences with only prosodic and broad phonotactic cues, Exp. 3 with only rhythmic cues, and Exp. 4 with rhythmic and broad phonotactic cues.

Experiment 1: Natural speech

This first experiment aims to test the discrimination of two languages in the most unconstrained condition, using natural, unsynthesized sentences. The two languages we have selected are Dutch and Japanese. The discrimination of this pair of languages was previously tested in 2-3 month-old English infants, and yielded only a marginally significant result (Christophe & Morton, 1998). This was interpreted as showing a growing focus on the native language, hence a loss of interest in foreign ones (consistent with Mehler et al., 1988). This pair of languages has never been tested on newborns, but it is expected to be easy to discriminate, given the English-Japanese discrimination by French newborns obtained by Nazzi et al. (1998), and the fact that English and Dutch are very similar in many respects, including rhythm.

Materials and Method

All the experiments included in this paper use the same methodology unless otherwise stated. We have made various attempts at improving certain aspects of the procedure; they are described in detail where appropriate.

---

[2] We also have unpublished data showing that newborns do not discriminate between Catalan and Spanish, both syllable-timed languages.

[3] The point of the present study is not, obviously, to deny that intonation can be processed by newborns and play a role in language acquisition (see for instance Guasti, Nespor, Christophe, & Ooyen, in press), but to ask whether rhythm is sufficient for babies to discriminate languages, and therefore whether they genuinely process rhythm.

Stimuli. Dutch and Japanese sentences were taken from a corpus constituted by Nazzi (1997; Nazzi et al., 1998), comprising short news-like sentences read by four female native speakers per language. This corpus consists exclusively of adult-directed speech. We selected 5 sentences per speaker, i.e., 20 sentences per language, matched in number of syllables (15 to 19, with an average of 17) and in duration (3120 ms ±186 for Dutch, 3040 ms ±292 for Japanese, $F(1,39) = 1.1$, $p = 0.3$). We were also concerned about the possibility that speakers in one language might have a higher pitch than speakers in the other language. Average fundamental frequency[4] is indeed significantly different between the two languages: 216 Hz ±19 for Dutch, 235 Hz ±15 for Japanese, $F(1,39) = 11.8$, $p = 0.001$. This is compensated for through resynthesis in Experiment 2, and we will see that this had no influence on discrimination. Sentences in subsequent experiments were resynthesized from these 40 source sentences, and differ only with respect to the type of synthesis that was used[5].

Experimental protocol. As is customary when testing newborns, we used the non-nutritive sucking technique in a habituation paradigm (Eimas, Siqueland, Jusczyk, & Vigorito, 1971). We have taken particular care to blind the experimenter with respect to the condition and to reduce direct interventions during the test to a minimum. For this purpose, the experiment has been programmed on a PC in such a way as to be maximally automatized.

Experimental conditions Babies are randomly assigned to the control or to the experimental group. In the habituation phase, they are exposed to 10 sentences uttered by two speakers of one language. In the test phase, babies from the control group hear 10 new sentences uttered by the other two speakers in the same language, whereas babies from the experimental group hear 10 new sentences uttered by two new speakers in the other language. The language presented in the habituation phase is counterbalanced across subjects, resulting in four experimental conditions. Assignment of the subjects to the different conditions is managed by the program and withheld from the experimenter.

Procedure The test takes place in a sound-attenuated booth, with only the experimenter and the baby inside. The experimenter sits out of the infant's field of vision and wears a sound-proof headset playing masking noise. This noise consists of four superimposed streams of all the experimental sentences playing continuously, in order to optimally mask the stimuli. The baby is seated in a reclining chair, and is presented with a pacifier fixed on a flexible arm. The air pressure in the pacifier is measured by a pressure transducer, amplified and transmitted to the computer via an analog/digital board. The computer detects sucks and computes their relative amplitude.

During the first two minutes, the baby sucks in silence and the computer calculates a high-amplitude (HA) threshold, such that 75% of the sucks have an amplitude above the threshold. Subsequently, only HA sucks are considered[6]. The habituation phase then starts. Each HA suck may elicit one sentence, but HA sucks produced while a sentence is already playing do not trigger an additional one, and two consecutive sentences are separated by at least 500 ms of silence. Sentences are played in a random order, directly from the hard disk of the computer, by two loudspeakers placed in front of the baby. The habituation phase goes on until the habituation criterion is met: it consists in a minimum 25% decrease in the number of HA sucks per minute for two consecutive minutes, compared with the maximum number of sucks previously produced in 1 minute[7]. When the criterion is met, the computer switches to the test phase, which lasts for 4 minutes. Test sentences are played in the same conditions as the habituation sentences.

Delay and rejection conditions Other factors may interfere with the test and may make it necessary to delay the shift to the test phase or simply to discard the baby's data. We have tried as much as possible to have these decisions made automatically by the computer, on the basis of the sucking pattern and of indications given by the experimenter. When the baby loses the pacifier, starts crying, or falls asleep, the experimenter needs to take appropriate action. When doing so he presses a special key on the keyboard, indicating the occurrence of an event interfering with the baby's sucking. The most critical period in the test consists of the two minutes before and the two after the shift, which are used for the statistical analyses. It is important to ensure that during those four minutes, (a) no interference has occurred, (b) the baby was awake and sucking, and thus heard enough sentences. The computer thus implements the following conditions:

Delay:

- if some interference was signaled during the 2 minutes preceding the shift,
- OR if the baby didn't hear any sentence during any one of those 2 minutes,

then the shift is delayed for at least one minute, and the habituation phase goes on.

Rejection:

- if some interference was signaled during the 2 minutes following the shift,

---

[4] Fundamental frequency was extracted at intervals of 5 ms using the Bliss software. We calculated an average F_0 for each sentence, as the average of all its non-zero F_0 values.

[5] Samples of the different types of stimuli used in the present experiments can be heard on: http://www.lscp.net/persons/ramus/resynth/ecoute.htm.

[6] Eliminating the weaker 25% of the sucks helps increasing the signal/noise ratio (Siqueland & DeLucia, 1969).

[7] The first minute of the phase is not taken into account for the determination of the maximum. Additional conditions impose that this maximum is at least 20 HA sucks per minute, and that the habituation phase lasts at least 5 minutes.

- OR if the baby didn't hear any sentence during any one of those 2 minutes,
- OR if the habituation phase has already lasted for 20 minutes,

then the test is discontinued and the baby's data are discarded.

In addition, the experimenter himself may make the decision to discontinue the test, (a) if the baby refuses the pacifier, (b) if he/she doesn't stay awake or suck enough, (c) if he/she keeps crying or being agitated.

Participants. Experiments took place at the Maternité Port-Royal, Hôpital Cochin in Paris. Participation was solicited from the mothers after birth, during their stay at the maternity hospital. Babies were pre-selected on the basis of their medical files, according to the following criteria:

- Age between 2 and 5 days old;
- Gestational age greater than or equal to 38 weeks;
- Birth weight greater than or equal to 2800 grams;
- No suffering at birth (APGAR score = 10 at 5 min);
- Normal medical assessments at birth and at two days;
- No seroconversion to rubella or toxoplasmosis;
- Mother not affected by viruses, and not addicted to any drug including alcohol or nicotine;
- No family history of deafness or neurological problems;
- No Dutch or Japanese spoken at home[8].

Babies were tested three hours after feeding on average, when they could be easily woken and kept in a quiet alert state, and when their sucking reflex was maximal.

In the present experiment, 32 babies were successfully tested, 18 males and 14 females, with a mean age of 67 ± 22 hours, a mean gestational age of $40 \pm 1;1$ weeks and a mean birth weight of 3530 ± 402 g. Twenty-nine came from monolingual French families, 2 from families where one or several other languages than French are spoken and 1 from a family where no French is spoken. The results of 42 additional babies were rejected for the following reasons: rejection of the pacifier (1), sleeping or insufficient sucking before the shift (12), crying or agitation (9), failure to meet the habituation criterion (9), sleeping or insufficient sucking after the shift (6), loss of the pacifier after the shift (4) and computer failure (1).

Results

Figure 1 shows the number of HA sucks per minute for the 2 groups of subjects. To ensure that babies were in comparable conditions during the habituation phase, an ANOVA was performed on average number of HA sucks over the 5 minutes preceding the shift, and showed no significant effect of group (control or experimental) [$F(1,31) = 2.6$, $p = 0.12$], although there is a trend for babies in the control group to suck more, and no significant effect of the language heard in habituation (Dutch or Japanese) [$F(1,31) < 1$]. In order to assess whether the experimental group reacted more to the change than the control group, we conducted an analysis of covariance (ANCOVA), with the average number of HA sucks during the 2 post-shift minutes as dependent variable, the average number of HA sucks during the 2 pre-shift minutes as covariate, and the group as independent variable[9]. Here, there is no significant group effect [$F(1,29) < 1$], showing that the babies in the experimental group have not reacted to the language change.

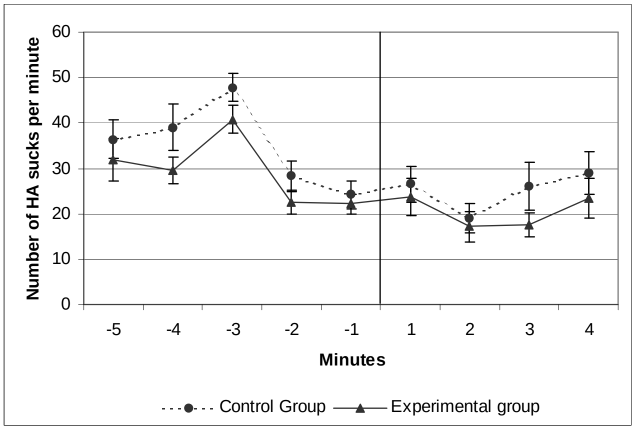

Figure 1. Exp. 1: Dutch-Japanese discrimination – Natural speech. Minutes are numbered from the shift, indicated by the vertical line. Error bars represent ±1 standard error of the mean. Adapted with permission from Ramus et al. (2000). Copyright 2000 American Association for the Advancement of Science.

Discussion

This experiment shows that newborns fail to discriminate between Dutch and Japanese when the stimuli are not degraded at all, consisting just of natural sentences. This may seem inconsistent with Nazzi et al.'s (1998) finding that similar French newborns can discriminate between English and Japanese. Among the few differences between our experiment and that of Nazzi et al., is the fact that their sentences were low-pass filtered, whereas ours aren't. While it may seem that the more information, the easier the discrimination, previous experiments on adults suggest that it is not always the case: Irrelevant information may actually impair the discrimination (Ramus & Mehler, 1999; Ramus, Dupoux, Zangl, & Mehler, submitted).

---

[8] Past studies show that newborns do not need to be familiar with one of the two languages to be able to discriminate them (Mehler & Christophe, 1995; Nazzi et al., 1998); familiarity only affects preference for one language (Mehler et al., 1988), which we are not testing here. Except for the two target languages, we therefore made no particular effort to discard babies from other linguistic backgrounds than French.

[9] This is the standard analysis of sucking rates since Christophe, Dupoux, Bertoncini, and Mehler (1994) showed that it is more appropriate than doing a simple ANOVA on a dishabituation index (e.g., the difference in sucking rates between the 2 minutes after and the 2 minutes before the shift). Here, we also ran the ANOVAs and found that they always led to the same conclusions as the ANCOVAs. We thus only report the results of the latter.

Here, the fact that each newborn hears 4 speakers during the experiment may constitute such irrelevant information. Indeed, it has been suggested that normalizing for speaker variability is a costly process in younger infants, that may interfere with other speech categorization abilities (Jusczyk, Pisoni, & Mullenix, 1992). It is actually remarkable that all language discrimination experiments performed on newborns to this day have used stimuli where speaker variability was either completely absent, through the use of a single bilingual speaker (Mehler et al., 1988; Moon et al., 1993), or at least strongly attenuated through the use of low-pass filtering (Nazzi et al., 1998). Experiment 1 is thus the first language discrimination experiment to expose newborns to 4 different voices.

Speech resynthesis is a convenient technique to test whether newborns were disturbed by speaker variability: whatever the original number of speakers, all sentences are synthesized using only one voice, that of the synthesizer. If our hypothesis is correct, we should then predict that newborns will be able to discriminate the two languages once the sentences are resynthesized.

Experiment 2: *Saltanaj* resynthesized speech

Materials and Method

Stimuli. We used the same sentences as for Experiment 1, but we resynthesized them in the *saltanaj* manner described by Ramus and Mehler (1999). For each sentence, the fundamental frequency (F_0) is measured, phonemes are manually identified and their duration measured, with the aid of speech analysis software. These parameters can be manipulated at will, and then used as input to the speech synthesizer MBROLA[10] (Dutoit, Pagel, Pierret, Bataille, & Vrecken, 1996). The *saltanaj* manipulation involves reducing the phonemic inventory of the sentences, by mapping each phoneme to a single instance of the same manner of articulation: fricatives are mapped to /s/, vowels to /a/, liquids to /l/, plosives to /t/, nasals to /n/ and glides to /j/. The exact phoneme durations as well the F_0 curve are copied from the original sentences. Thus, the overall rhythm and intonation of the sentences are faithfully preserved, together with broad phonotactic characteristics. However, phonetic and fine phonotactic differences are eliminated. Obviously, access to syntax and meaning is blocked as well. Voice differences are eliminated, since a single synthetic voice is used; however, prosodic differences between speakers are preserved. It therefore still makes sense to have a speaker change in the control condition, where "speakers" refers to those who produced the original sentences.

Incidentally, resynthesis gives full control over the average fundamental frequency which, we have noted earlier, was significantly different between the two languages. We have thus decided to reduce this difference by multiplying all F_0 values by 1.04 for Dutch, and by 0.96 for Japanese. Note that although this manipulation removes a possible confound, it is not a very plausible one. Indeed, in Experiment 1, the fact that Japanese speakers have a higher pitch on average isn't sufficient by itself to allow for a discrimination.

Procedure. The procedure was the same as for Exp. 1, except that we tried to adapt the 2-shift design of Hesketh, Christophe, and Dehaene-Lambertz (1997) to experimentation on newborns. The first two phases of the experiment were run exactly as in Exp. 1. After a baby had undergone the first shift and the four-minute test phase, a second shift was made possible. It was subject to the same habituation criterion, delay and rejection conditions as the first one. When met, it allowed the baby to undergo a second four-minute test phase. In that case, the second shift was of a different kind (language or speaker) to the first one. Babies who didn't successfully pass the second shift were nevertheless kept for analysis of the first shift. Thus, this attempt at improving the procedure did not interfere with the collection of the data on the first shift, making it possible to independently analyze the results of the two shifts. Indeed, for all babies, everything was as in Exp. 1 up until the fourth test minute. Some babies were just allowed to go on further if they could. This experiment was stopped as soon as 32 babies successfully passed the first shift.

Participants. Thirty-two babies were successfully tested, 16 males and 16 females, with a mean age of 80 ± 25 hours, a mean gestational age of $39 \pm 0;6$ weeks and a mean birth weight of 3508 ± 477 g. Twenty-five came from monolingual French families, 3 from families where one or several other languages than French are spoken and 4 from a family where no French is spoken. The results of 20 additional babies were rejected for the following reasons: sleeping or insufficient sucking before the first shift (6), crying or agitation (4), failure to meet the habituation first criterion (1), sleeping or insufficient sucking after the first shift (3), loss of the pacifier after the first shift (6).

Results

Figure 2 shows the number of HA sucks per minute for the two groups of babies. There was no significant group effect on babies' sucking during the 5 pre-shift minutes [$F(1,31) < 1$], neither was there an effect of the habituation language [$F(1,31) = 1.8$, $p = 0.19$]. An ANCOVA on the 2 post-shift minutes, controlling for the 2 pre-shift minutes, showed a significant group effect [$F(1,29) = 6.3$, $p = 0.018$], indicating that babies in the experimental group increased their sucking significantly more than those in the control group. This leads us to the conclusion that the babies in the experimental group were able to discriminate between the two languages.[11]

---

[10] MBROLA is freely available from http://tcts.fpms.ac.be/synthesis/mbrola.html.

[11] Out of 32 babies successfully passing the first shift, only 11 passed the second one. The results of the remaining 21 were discarded due to sleeping or insufficient sucking before the second shift (7), crying or agitation (2), failure to meet the second habituation criterion before the 20 minute time limit (6), sleeping or in-

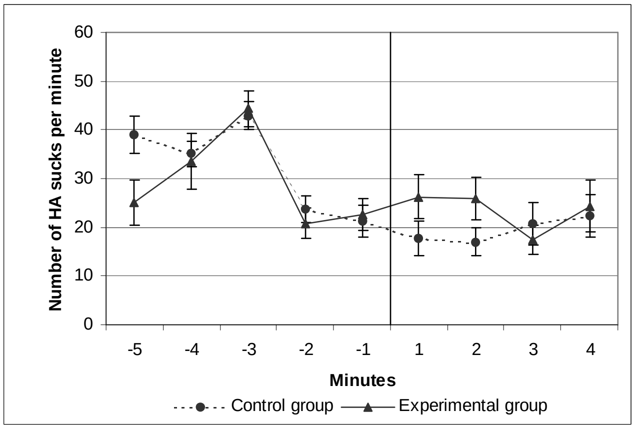

Figure 2. Exp. 2: Dutch-Japanese discrimination – *Saltanaj* speech, first shift. Minutes are numbered from the shift, indicated by the vertical line. Error bars represent ±1 standard error of the mean. Adapted with permission from Ramus et al. (2000). Copyright 2000 American Association for the Advancement of Science.

Discussion

The data obtained on the 32 newborns who successfully passed the first shift show that (a) they are able to discriminate between Dutch and Japanese, (b) they can do so when sentences are resynthesized in the *saltanaj* manner, i.e. when lexical, syntactic, phonetic and some phonotactic information is removed.

Although the interaction with Exp. 1 is not quite significant $[F(1,59) = 2.6, p = 0.11]$[12], this is also consistent with the hypothesis that newborns have difficulties coping with talker variability (Jusczyk et al., 1992), which would be the reason why they failed to discriminate the same sentences when they were not resynthesized.

The *saltanaj* resynthesis achieves a comparable level of stimulus degradation as low-pass filtering: Since all the durations and the fundamental frequency are faithfully reproduced, prosody, in a broad sense, is still preserved. It is therefore insufficient to disentangle the role of rhythm and intonation. This concern is addressed in the next two experiments.

Experiment 3: *sasasa* with artificial intonation

Materials and Method

Stimuli. In previous experiments testing language discrimination by adults on the basis of rhythm only (Ramus & Mehler, 1999; Ramus et al., submitted), sentences were resynthesized in the *flat sasasa* manner: all consonants were mapped to /s/ and all vowels to /a/, and in addition the original F_0 contour of the sentence was ignored and replaced by a constant F_0. Thus all differences concerning intonation or syllable structure were eliminated, preserving only rhythmical differences between the two languages.

When testing babies, an additional concern is to keep them awake and active in the experiment. In this respect, *flat sasasa* stimuli are potentially problematic. Both their low phonetic diversity and their monotonous intonation are susceptible to provoke boredom or distress in infants, and/or to induce them to process the stimuli as non-speech. We thus felt we had to improve the attractiveness of our stimuli, while still adequately testing our hypotheses. Considering that newborns are known to react normally to low-pass filtered speech (Mehler et al., 1988; Nazzi et al., 1998), we assumed that phonetic diversity was not a necessary condition, but we chose to preserve some variability in the intonation.

Therefore, we decided to resynthesize the same 40 sentences as before using a *sasasa* phonetic mapping, i.e., to map each consonant to /s/ and each vowel as /a/. However, instead of applying a flat intonation to each sentence, we applied artificial intonation contours. Five intonation contours inspired from French sentences were designed and each was applied to 4 of the Dutch and 4 of the Japanese sentences. All contours included a regular declination towards their end, in order to be more easily adapted to sentences of different lengths; they are illustrated in Figure 3. Thus, the resynthesized sentences incorporate both within-sentence and within-language intonational variability, but no differences between the two languages.

A potential criticism of this method is that there might be an interaction between intonation and rhythmic structure, so that the five contours selected might be more adapted to the rhythmic structure, say, of Dutch, than to that of Japanese. This would then introduce a difference between the two sets of sentences that would not be a rhythmic difference, strictly speaking. In order to investigate this possibility, we conducted the following preliminary experiment on adult subjects.

Judgement by adult subjects of sentences resynthesized with an artificial intonation Twelve participants were recruited and tested in a quiet room. They were 4 men and 8 women, with a mean age of 34 years, and of various native languages (4 French, 1 Rumanian, 3 Spanish, 1 German, 2 English, 1 Dutch).

Two blocks of stimuli were designed: one comprising the 40 experimental sentences to be used on the babies (*sasasa*

---

sufficient sucking after the second shift (4), loss of the pacifier after the second shift (2). The small proportion of babies able to undergo the second shift already shows that this procedure, as used on 2-month-olds, is not viable for newborns: it would lead to discard too much data (here, a total of 41 babies out of 52). In addition, a discrimination index computed as in Hesketh et al., 1997 was not significantly different from 0 $[t(10) < 1]$, indicating that these 11 babies did not increase their sucking more after the experimental shift than after the control one. Rather, they tended to persevere in the behavior elicited by the first shift. We conclude that there is little to learn from a second shift with newborns.

[12] Note that interactions are seldom significant in experiments on newborns anyway, due to their low statistical power. For instance, in directly comparable studies, no significant interaction were ever reported (Mehler et al., 1988; Nazzi et al., 1998).

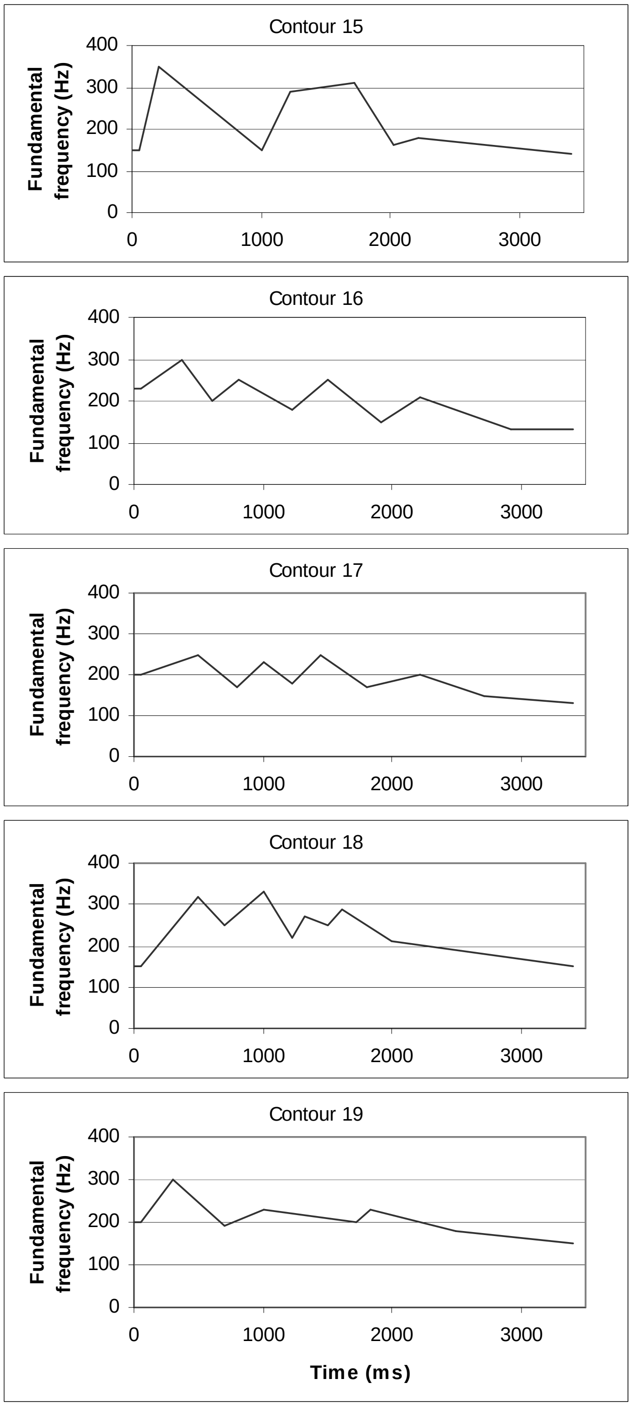

Figure 3. Intonation contours used in Exp. 3.

with artificial intonation), and the other the *saltanaj* sentences with original intonation used in Exp. 2, to provide a baseline.

Each participant heard the sentences one by one, in a random order within each block; the order of the blocks was counterbalanced across subjects. The task was to judge how natural the intonation of each sentence was (from 0: very strange to 5: perfectly natural). If artificial intonations are equally adapted to the rhythmic structure of the two languages, they should yield similar average judgements. These are shown in Figure 4.

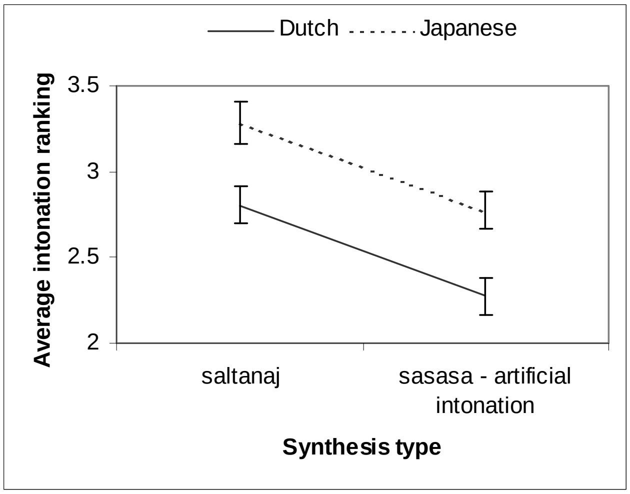

Figure 4. Adult subjects' judgements of the intonation pattern of Dutch and Japanese sentences. Error bars represent ±1 standard error of the mean by subject.

It appears that the artificial intonations of the *sasasa* stimuli are judged to be significantly less natural in Dutch than in Japanese sentences [$F(1,11) = 12.5$, $p = 0.005$]. However, the same is true of the *saltanaj* sentences with their original intonation [$F(1,11) = 8.4$, $p = 0.02$]. It thus can't be interpreted as an effect of mismatch between the artificial intonation contours and Dutch rhythm. Rather, it seems to reflect the influence of syllabic structure on subjects' responses, although they were instructed to report specifically about intonation. From their reports, it appears that the presence of heavy consonant clusters in Dutch (also reflected by longer /s/s in the *sasasa* version) biased them in favor of Japanese. Thus, there is a main effect of language [$F(1,11) = 21.8$, $p = 0.001$], and there is also a main effect of type of synthesis [$F(1,11) = 8.4$, $p = 0.02$], revealing that *sasasa* synthesis sounded less natural to the subjects than *saltanaj*. Despite these effects that interfered with the subjects' judgements, it is most important to note that there is no interaction between language and type of stimuli [$F(1,11) < 1$], indicating that the artificiality of the *sasasa* stimuli's intonation did not interact with the rhythmic structure of the two languages[13], which was the hypothesis under test. We therefore consider

---

[13] Note also that no floor nor ceiling effect would have prevented this interaction to emerge.

our stimuli as appropriate to test language discrimination on the basis of rhythm only.

Procedure. Due to the unsuccessful attempt to have the babies undergo 2 shifts in Exp. 2, we abandoned the second shift and returned to the one-shift procedure used in Exp. 1.

Participants. Forty babies[14] were successfully tested, 19 males and 21 females, with a mean age of 66 ± 23 hours, a mean gestational age of $40;1 \pm 1;1$ weeks and a mean birth weight of 3428 ± 424 g. Twenty-six came from monolingual French families, 11 from families where one or several other languages than French are spoken and 3 from families where no French is spoken. The results of 20 additional babies were rejected for the following reasons: rejection of the pacifier (2), sleeping or insufficient sucking before the shift (7), crying or agitation (3), failure to meet the habituation criterion (1), sleeping or insufficient sucking after the shift (4), loss of the pacifier after the shift (3).

Results

Figure 5 shows the number of HA sucks per minute for the two groups of babies. There was no significant group effect on babies' sucking during the 5 pre-shift minutes [$F(1,39) = 1.5$, $p = 0.23$]. However, there was a significant effect of the habituation language [$F(1,39) = 12.8$, $p = 0.001$], babies listening to Dutch sucking more (35.6 ± 8.3 sucks per minute on average) than those listening to Japanese (25.8 ± 9.1 sucks per minute). To take this effect into account, we included the habituation language factor in the usual ANCOVA. We found that it had no significant effect on sucking patterns during the 2 post-shift minutes [$F(1,35) < 1$], and most importantly, that it did not interact with the group factor [$F(1,35) < 1$]. Thus, this effect had no consequence on the overall result of the experiment; its interpretation will be addressed in a later section. Finally, the ANCOVA on the 2 post-shift minutes, controlling for the 2 pre-shift minutes, showed no group effect [$F(1,35) < 1$], indicating that the babies didn't discriminate between the two languages.

Discussion

It is worth noting that, although the sucking patterns around the shift of language unambiguously show an absence of dishabituation to the new language, the fact that babies sucked significantly more to listen to Dutch than to Japanese in the habituation phase yet suggests that the two languages are not entirely the same to them. This might, of course, be a sampling effect, i.e., babies who intrinsically suck more being assigned by chance to the "Dutch first" condition; yet, both the size and the significance of the effect make this interpretation unlikely. We now leave this issue aside to return to it in a later section.

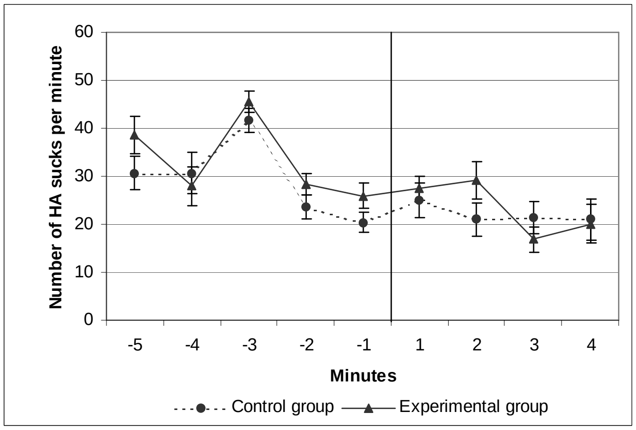

Figure 5. Exp. 3: Dutch-Japanese discrimination – *Sasasa* speech with artificial intonation. Minutes are numbered from the shift, indicated by the vertical line. Error bars represent ±1 standard error of the mean.

There are several possible interpretations of the failure of babies to discriminate between Dutch and Japanese given the *sasasa* sentences with artificial intonation. One is that babies don't process speech rhythm, and that language discrimination experiments should be re-interpreted as revealing the processing of intonation. Another is that rhythm and intonation are inseparable: babies may be sensitive to speech rhythm, but intonation is necessary to fully process it. Yet another interpretation would be that babies can process speech rhythm, but the stimuli used were inadequate for them to exhibit this ability.

For instance, Ramus et al. (2000) showed that newborns don't discriminate Dutch from Japanese anymore when the same sentences are played backwards. This suggests that stimuli that are not enough speech-like are not correctly processed by babies, even though they contain enough basic acoustic information for the discrimination to be feasible in principle. In this respect, *sasasa* might not be speech-like enough: there is indeed no natural language with so little phonetic diversity. It could also be that these stimuli are too boring or distressing for the babies, as we had hypothesized regarding a flat intonation. Whatever the appropriate explanation, we will now try to increase the chances that the babies correctly process the stimuli.

Experiment 4: *saltanaj* with artificial intonation

Materials and Method

Stimuli. There are two differences between the stimuli used in Experiment 2 and those of Experiment 3: one is the reduction of the phonetic inventory (from *saltanaj* to *sasasa*), and the other is the use of artificial intonation contours instead of the original ones. At least one of them has caused babies to fail in the discrimination task. It is therefore natural to undo one of those changes in order to know which was critical. We thus returned to the *saltanaj* stimuli of Exp. 2, but this time we applied them the artificial intonation contours of Exp. 3.

---

[14] Eight additional babies were tested after a first analysis on the first 32 babies, because at that stage it was not clear if there was a trend that was not significant by lack of power, or no effect at all.

Participants. Forty babies were successfully tested, 21 males and 19 females, with a mean age of 68 ± 21 hours, a mean gestational age of $40;1 \pm 1$ weeks and a mean birth weight of 3512 ± 341 g. Twenty-seven came from monolingual French families, 12 from families where one or several other languages than French are spoken and one from a family where no French is spoken. The results of 44 additional babies were rejected for the following reasons: rejection of the pacifier (5), sleeping or insufficient sucking before the shift (22), crying or agitation (8), failure to meet the habituation criterion (2), sleeping or insufficient sucking after the shift (4), loss of the pacifier after the shift (3).

Results

Figure 6 shows the number of HA sucks per minute for the two groups of babies. There was no significant group effect on babies' sucking during the 5 pre-shift minutes [$F(1,39) < 1$], neither was there an effect of the habituation language [$F(1,39) = 2.1$, $p = 0.16$]. An ANCOVA on the 2 post-shift minutes, controlling for the 2 pre-shift minutes, shows no significant group effect [$F(1,37) = 1.46$, $p = 0.24$]. However, examination of Figure 6 suggests that there is an effect, which is confined to the first minute after the shift. A new ANCOVA, taking as dependent variable the number of sucks during the first post-shift minute, and controlling for the 2 pre-shift minutes, yields a significant group effect indeed [$F(1,37) = 4.48$, $p = 0.04$]. This suggests that the newborns have again discriminated between Dutch and Japanese. However, the effect is weaker than in Experiment 2, being evident during only one minute following the language change.

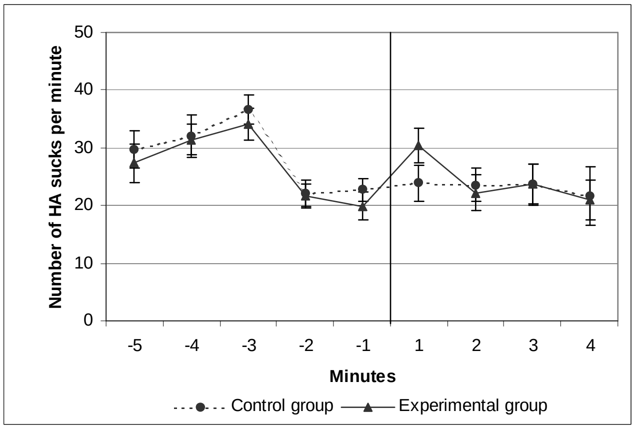

Figure 6. Exp. 4: Dutch-Japanese discrimination – *Saltanaj* speech with artificial intonation. Minutes are numbered from the shift, indicated by the vertical line. Error bars represent ±1 standard error of the mean.

Discussion

Obviously, this last experiment would need to be replicated. From a methodological point of view, it is not very satisfying to first look where the effect is located (here, on the first minute post-shift), and then restrict the analysis accordingly, since this increases the risk of a Type I error. However, although analyses based on two minutes before and two minutes after the shift have emerged as the methodological standards in the literature, deviations from this norm are by no means exceptional: analyses restricted to the first post-shift minute (e.g., McAdams & Bertoncini, 1997) or extended to the three post-shift minutes (e.g., Nazzi et al., 1998, Exp. 3) sometimes appear necessary to establish an effect. Awaiting replication, we will assume that the present discrimination is reliable for the remainder of the discussion.

Since no intonational difference remained between the two languages in the present *saltanaj* stimuli, intonation is not likely to be the cue whose processing is responsible for the language discrimination patterns observed. Here, the only cues available for discrimination were the sentences' rhythm and their broad phonotactic patterns. This leaves us with two possible interpretations.

Considering that newborns did not discriminate Dutch from Japanese in Exp. 3, when only rhythm was available, but did discriminate when some phonotactic information was added to rhythm in Exp. 4, the most straightforward interpretation is that newborns actually discriminated between the respective phonotactic patterns of Dutch and Japanese, e.g., they noticed the fact that there are many consonant clusters in Dutch but not in Japanese. However, such an interpretation is at odds with quite a large body of evidence. Indeed, sensitivity to phonotactic differences has been directly tested in experiments where newborns had to discriminate between lists of words of different syllabic structure. For instance, newborns were unable to discriminate bi-syllabic words with complex syllabic structure (e.g., CVCCCV, CCVCCV, CVCCVC...) from bi-syllabic words with simple syllabic structure (e.g., CVCV, VCCV, VCVC...), although they were able to discriminate simple bi-syllabic (CVCV) from tri-syllabic (CVCVCV) words (Bijeljac-Babic, Bertoncini, & Mehler, 1993). Similarly, they were unable to discriminate between tri-moraic (CVCCV) and bi-moraic words (CVCV), although they were again able to discriminate bi-syllabic from tri-syllabic words (Bertoncini, Floccia, Nazzi, & Mehler, 1995). If newborns were able to extract phonotactic regularities from 3-second long Dutch and Japanese utterances, they would be expected to do so also from bi-syllabic words. The fact that they do not suggests that sensitivity to phonotactic differences is not available at birth; this is also consistent with evidence that familiarity with the native language's phonotactic pattern emerges between 6 and 9 months of age (Jusczyk, Cutler, & Redanz, 1993; Friederici & Wessels, 1993; Jusczyk, Luce, & Charles-Luce, 1994), while familiarity with the native language's prosodic pattern seems to emerge soon after birth (Mehler et al., 1988; Christophe & Morton, 1998).

The other possible interpretation of our results is that newborns perceived the rhythmical differences between the two languages. This interpretation assumes that the reason for the failure to discriminate in Exp. 3 may lie in some aspect of the *sasasa* stimuli that prevented the babies from correctly processing them. For instance, their extremely low phonetic diversity might lead babies to process them as non-speech. Alternatively, the constant frication of *sasasa* sentences may have been too distressing for babies to correctly perform the task; indeed, adult subjects also rated these stimuli lower than the *saltanaj* in Exp. 3, and complained about the harsh sound of the /s/s. Whatever the correct explanation may be, the alternatives could be tested by running further discrimination experiments while manipulating the nature and the variety of the phonemes used in the resynthesis. Possibly *mamama* stimuli would be preferable to *sasasa*. One could also resynthesize the sentences as *salatanaja*, a transformation similar to the *sasasa* in that consonant clusters would be mapped to a single phoneme of the same total duration, but the nature of this phoneme would be allowed to vary randomly. This will be matter for future investigations.

Post-hoc analysis: Increased sucking for Dutch

After finding a significant effect of habituation language in Experiment 3 (with babies sucking more to Dutch than to Japanese), we have looked for a similar trend in the other experiments. As shown in Table 1 and Figure 7, the same trend was present in all four experiments, suggesting a consistent phenomenon.

Table 1
Average number of HA sucks produced during the 5 pre-shift minutes, as a function of Experiment and language.

	Exp. 1	Exp. 2	Exp. 3	Exp. 4
Dutch	33 ± 11	32.2 ± 11.2	35.6 ± 8.3	29.4 ± 9.6
Japanese	30.9 ± 11.1	27 ± 10.4	25.8 ± 9.1	25.3 ± 8.3

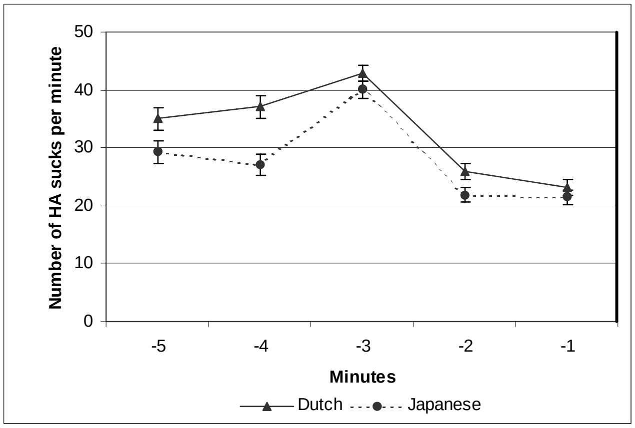

Figure 7. Average number of HA sucks produced during each of the 5 pre-shift minutes, as a function of language. Data of Experiments 1-4 are collapsed. Error bars represent ±1 standard error of the mean.

Overall, during the 5 minutes preceding the shift, babies produced on average 32.5 ± 10 HA sucks per minute to listen to Dutch, and 27.1 ± 9.7 to listen to Japanese. The difference is significant: $F(1,142) = 11.1$, $p = 0.001$. The fact that this effect is consistent across 4 experiments suggest that this is not a random sampling effect. It is remarkable that a similar effect was also noted by Nazzi et al. (1998): they found that French newborns sucked more to listen to English than to listen to Japanese. In earlier studies, such patterns had been interpreted as showing babies' "preference" for or at least familiarity with their native-language: Mehler et al. (1988) found that French newborns sucked more during the habituation phase to listen to French than to Russian, and Moon et al. (1993), using a technique directly assessing preference, found that newborns sucked more to listen to their native language (English or Spanish). Similar results have been found using preferential looking procedures in older babies (Dehaene-Lambertz & Houston, 1998; Bosch & Sebastián-Gallés, 1997).

Here (and also in Nazzi et al.'s study), where neither language was the babies' native language, one may want to look for an alternative explanation. For instance, following adult subjects' judgements in Exp. 3, one might argue that Dutch sentences, containing more consonant clusters and more frication, are less pleasant to listen to, and thus keep the babies more excited. Such an interpretation would, however, assume that stimuli that are less preferred provoke more sucking, which would appear to be in contradiction with the earlier studies' interpretations. In the absence of a good model of what provokes a baby's sucks, the question remains open[15].

It is yet possible to provide an interpretation of the present results in terms of genuine preference or familiarity. Indeed, French can be seen as closer to Dutch and English than to Japanese along a number of dimensions. Regarding the most relevant one, rhythm, objective acoustic/phonetic measures of rhythmic properties suggest that French rhythm is much closer to that of Dutch and English than to that of Japanese (Ramus et al., 1999). Similar arguments could be made for syllable structure, intonation, size of the phonemic inventory... It is thus conceivable that Dutch and English sound more familiar to the French newborn than Japanese[16]. Native-language preference might therefore be re-interpreted as preference for the most familiar stimulus (along the dimensions that are relevant to the baby).

Obviously, these interpretations are very tentative and should not be taken as strong claims. If one really wants to test whether French newborns have a preference for Dutch or English versus Japanese, then one should use a procedure specially designed to assess preference, not discrimination. The fact remains that the trend is present in every single of five experiments to date, and awaits an explanation. This also suggests that when this effect reaches signifi-

---

[15] Note that this problem is not particular to the sucking behavior. In studies using preferential listening techniques, preference sometimes goes for the novel stimulus, and sometimes for the familiar one. No generalized account of infants' preferences has been proposed.

[16] Although newborns' native language was not strictly controlled in the present experiments, a large majority of babies had French as the main language in their family, and none had Dutch or Japanese.

cance ($p = 0.001$) in Exp. 3, it can hardly be dismissed as a sampling effect. However, we will not go as far as to claim that this differential processing of the two languages in the habituation phase amounts to discrimination, when a direct measure of discrimination provides no evidence thereof. At this point, we have to admit that no framework is available to interpret such effects.

Conclusion

The literature on infant speech perception has suggested for years that newborns discriminate languages' rhythmic patterns. The evidence accumulated so far, although compelling, has always left open the possibility that the discriminations observed might be due to a sensitivity to intonational differences. Here, using resynthesized stimuli, we have dissociated intonation from rhythm. The results suggest that intonation was not necessary for newborns to discriminate between Dutch and Japanese (Exp. 4), and therefore that rhythmic differences between these two languages have been processed by newborns. However, this interpretation is limited in two ways.

First, newborns appeared to discriminate the two languages only when some phonotactic information was present (*saltanaj*, Exp. 4), not when purely rhythmic cues were isolated (*sasasa*, Exp. 3). As we discussed earlier, there is relatively convincing previous evidence that sensitivity to phonotactic differences is not available at such an early age. The interpretation we favor therefore appeals to the inappropriateness of the *sasasa* stimuli for babies. Of course, this hypothesis would itself need to be bolstered by further experiments using more appropriate, yet purely rhythmic, stimuli. This will be matter for future research.

Another problem is that, even with the phonotactic cues present in Exp. 4, the discrimination effect found was relatively weak. This obviously invites an attempt to replicate. We may be luckier and find a stronger effect than the one presented here. Alternatively, we might also confirm that pure rhythmic differences do not elicit such reliable discriminations as differences along several dimensions. It is indeed possible that, when dissociated from intonation, rhythm loses part of its salience. After all, in real speech, rhythm and intonation are highly correlated: for instance, stress is signalled in terms of both the duration and the pitch of the syllable (and its energy); similarly, phonological phrase boundaries are marked both by final syllable lengthening and pitch movements. Thus, intonation partly underlines the information provided by purely durational cues. This means that, in addition to providing specific intonational cues, it introduces redundant rhythmic cues that may help rhythm processing itself. Indeed, it has been suggested that integrating multiple correlated cues might be a more powerful learning strategy than just considering isolated cues (Shi, Morgan, & Allopenna, 1998; Christiansen, Allen, & Seidenberg, 1998). Therefore, the most conservative interpretation of the present set of results at this stage may be that the integration of intonational and rhythmic cues better serves language discrimination than either of these cues taken separately. One could further test this interpretation by presenting resynthesised sentences where intonation would be preserved, but rhythm would be made the same for the two languages: one would predict a similar pattern of results as in Exp. 4, i.e., a weaker effect than in Exp. 2. Another prediction would be that when rhythmic and intonational differences are both present, but de-correlated (e.g., by permuting intonation contours across sentences within each language), language discrimination would be similarly impaired, as compared to the condition where the cues are correlated.

In conclusion, a definitive interpretation of the present set of experiments will have to await future results. Rhythm is still a good candidate as a perceptual cornerstone for the baby. But it may be futile to try to isolate the effects of purely rhythmic regularities, as their integration with other correlated cues may be the key to a successful processing of the speech signal.

References

Abercrombie, D. (1967). *Elements of general phonetics.* Chicago: Aldine.

Bertoncini, J., Floccia, C., Nazzi, T., & Mehler, J. (1995). Morae and syllables: Rhythmical basis of speech representations in neonates. *Language and Speech*, *38*, 311-329.

Bijeljac-Babic, R., Bertoncini, J., & Mehler, J. (1993). How do four-day-old infants categorize multisyllabic utterances? *Developmental Psychology*, *29*, 711-721.

Blevins, J. (1995). The syllable in phonological theory. In J. A. Goldsmith (Ed.), *The handbook of phonological theory* (p. 206-244). Cambridge: Blackwell.

Bosch, L., & Sebastián-Gallés, N. (1997). Native language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, *65*, 33-69.

Christiansen, M., Allen, J., & Seidenberg, M. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes*, *13*, 221-268.

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*(3), 1570-1580.

Christophe, A., Guasti, T., Nespor, M., Dupoux, E., & Ooyen, B. van. (1997). Reflections on phonological bootstrapping: Its role for lexical and syntactic acquisition. *Language and Cognitive Processes*, *12*(5/6), 585-612.

Christophe, A., & Morton, J. (1998). Is Dutch native English? Linguistic analysis by 2-month-olds. *Developmental Science*, *1*(2), 215-219.

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* (p. 87-99). Mahwah, NJ: Lawrence Erlbaum Associates.

Cutler, A., & Mehler, J. (1993). The periodicity bias. *Journal of Phonetics*, *21*, 103-108.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, *11*, 51-62.

Dauer, R. M. (1987). Phonetic and phonological components of language rhythm. In *XIth International Congress of Phonetic Sciences* (Vol. 5, p. 447-450). Tallinn.

Dehaene-Lambertz, G., & Houston, D. (1998). Faster orientation latencies toward native language in two-month old infants. *Language and Speech*, *41*(1), 21-43.

Demany, L., McKenzie, B., & Vurpillot, E. (1977). Rhythm perception in early infancy. *Nature*, *266*(5604), 718-9.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1568-1578.

Dupoux, E., Pallier, C., Kakehi, K., & Mehler, J. (2001). New evidence for prelexical phonological processing in word recognition. *Language and Cognitive Processes*, *16*(5/6), 491-505.

Dutoit, T., Pagel, V., Pierret, N., Bataille, F., & Vrecken, O. van der. (1996). The MBROLA Project: Towards a set of high-quality speech synthesizers free of use for non-commercial purposes. In *ICSLP'96.* Philadelphia. (MBROLA is freely available from http://tcts.fpms.ac.be/synthesis/mbrola.html)

Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*, 303-306.

Fowler, C., Smith, M., & Tassinary, L. (1986). Perception of syllable timing by prebabbling infants. *Journal of the Acoustical Society of America*, *79*(3), 814-825.

Friederici, A., & Wessels, J. (1993). Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception & Psychophysics*, *54*(3), 287-295.

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* (p. 3-48). Cambridge UK: Cambridge University Press.

Guasti, M. T., Nespor, M., Christophe, A., & Ooyen, B. van. (in press). Pre-lexical setting of the head-complement parameter through prosody. In J. Weissenborn & B. Höhle (Eds.), *How to get into language: Approaches to bootstrapping in early language development.* Amsterdam: Benjamins.

Hesketh, S., Christophe, A., & Dehaene-Lambertz, G. (1997). Nonnutritive sucking and sentence processing. *Infant Behavior and Development*, *20*, 263-269.

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

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

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*(3/4), 159-207.

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.

Jusczyk, P. W., Pisoni, D. B., & Mullenix, J. (1992). Some consequences of stimulus variability on speech processing by 2-month-old infants. *Cognition*, *43*, 253-291.

Ladefoged, P. (1975). *A course in phonetics.* New York: Harcourt Brace Jovanovich.

Low, E. L., Grabe, E., & Nolan, F. (2000). Quantitative characterisations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech*, *43*(4), 377-401.

Maidment, J. A. (1976). Voice fundamental frequency characteristics as language differentiators. *Speech and hearing: Work in progress, University College London*, 74-93.

Maidment, J. A. (1983). Language recognition and prosody: further evidence. *Speech, hearing and language: Work in progress, University College London*, *1*, 133-141.

McAdams, S., & Bertoncini, J. (1997). Organization and discrimination of repeating sound sequences by newborn infants. *Journal Of The Acoustical Society Of America*, *102*(5), 2945-53.

Mehler, J., & Christophe, A. (1995). Maturation and learning of language during the first year of life. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (p. 943-954). Bradford Books / MIT Press.

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

Mehler, J., Sebastián-Gallés, N., Altmann, G., Dupoux, E., Christophe, A., & Pallier, C. (1993). Understanding compressed sentences: the role of rhythm and meaning. *Annals of the New York Academy of Science*, *682*, 272-282.

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, *16*, 495-500.

Morgan, J. L., & Demuth, K. (1996). *Signal to Syntax: Bootstrapping from speech to grammar in early acquisition.* Mahwah NJ: Lawrence Erlbaum Associates.

Morton, J., Marcus, S., & Frankish, C. (1976). Perceptual centers (P-centers). *Psychological Review*, *83*(5), 405-408.

Nazzi, T. (1997). *Du rythme dans l'acquisition et le traitement de la parole.* doctoral dissertation, Ecole des Hautes Etudes en Sciences Sociales.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 756-766.

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

Pallier, C., Sebastian-Galles, N., Dupoux, E., Christophe, A., & Mehler, J. (1998). Perceptual adjustment to time-compressed speech: a cross-linguistic study. *Memory & Cognition*, *26*, 844-851.

Pijper, J. R. de. (1983). *Modelling British English intonation.* Dordrecht - Holland: Foris.

Pike, K. L. (1945). *The intonation of American English.* Ann Arbor, Michigan: University of Michigan Press.

Pinker, S. (1984). *Language learnability and language development.* Cambridge: Harvard University Press.

Prince, A., & Smolensky, P. (1993). *Optimality theory: Constraint interaction in generative grammar* (Tech. Rep. No. TR-2). Rutgers University.

Ramus, F., Dupoux, E., Zangl, R., & Mehler, J. (submitted). An empirical study of the perception of language rhythm.

Ramus, F., Hauser, M., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, *288*(5464), 349-351.

Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *Journal of the Acoustical Society of America*, *105*(1), 512-521.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*(3), 265-292.

Roach, P. (1982). On the distinction between "stress-timed" and "syllable-timed" languages. In D. Crystal (Ed.), *Linguistic controversies.* London: Edward Arnold.

Sebastián-Gallés, N., Dupoux, E., & Costa, A. (2000). Adaptation to time-compressed speech: Phonological determinants. *Perception & Psychophysics*, *4*, 834-842.

Shi, R., Morgan, J. L., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: a cross-linguistic perspective. *Journal of Child Language*, *25*, 169-201.

Siqueland, E. R., & DeLucia, C. A. (1969). Visual reinforcement of nonnutritive sucking in human infants. *Science*, *165*(898), 1144-6.