# DISCRIMINATION OF PITCH CONTOURS BY NEONATES

**Thierry Nazzi**
**Caroline Floccia**
**Josiane Bertoncini**
*Laboratoire de Sciences Cognitives et Psycholinguistique*

With the high-amplitude sucking procedure, newborns were presented with two lists of phonetically varied Japanese words differing in pitch contour. Discrimination of the lists was found, thus indicating that newborns are able to extract pitch contour information at the word level.

**infant speech perception pitch neonates**

From the acoustical point of view, prosody usually refers to dimensions such as the duration, fundamental frequency or pitch contour, and intensity of the elements that constitute the speech wave. The close, although not perfect, correspondence between prosody and structural syntactic properties has prompted the proposal of a theoretical framework, called the *prosodic bootstrapping hypothesis* (Gleitman & Wanner, 1982; Morgan & Demuth, 1996). This hypothesis states that processing of prosodic information facilitates the acquisition of some structural properties of the language.

The proposal of the prosodic bootstrapping hypothesis has enhanced interest in the study of young infants' sensitivity to prosodic properties. For example, regarding prosodic information at the level of utterances and sentences, Fernald & Kuhl (1987) showed that intonational and rhythmic information determines young infants' preference for infant-directed speech over adult-directed speech. In addition, Nazzi, Bertoncini & Mehler (1998) found that rhythmic properties might play a crucial role in newborns' ability to discriminate between languages.

Infants' sensitivity to prosodic information at the word level has also been investigated. Infants discriminated two words whose final vowels differed either in duration (Eilers, Bull,

• **Thierry Nazzi,** Institute of Human Development, University of California at Berkeley, 1203 Tolman Hall, MC 1690, Berkeley, CA 94720-1690; Tel: (510) 643-9739; Fax: (510) 642-7969; e-mail: thierry@1scp.ehess.fr.

INFANT BEHAVIOR & DEVELOPMENT 21 (4), 1998, pp. 779–784
ISSN 0163-6383

Oller & Lewis, 1984), amplitude (Bull, Eilers & Oller, 1984), mean fundamental frequency (Bull, Eilers & Oller, 1985) and fundamental frequency contour (Karzon & Nicholas, 1989). Infants' processing of prosodic information may also interact with their processing of other linguistic dimensions. For example, two-month-old infants discriminated a consonant contrast in both stressed and unstressed syllables of bisyllabic words (Jusczyk & Thompson, 1978), but discriminated a similar consonant contrast only if it appeared in the stressed syllable of trisyllabic words (Karzon, 1985). All of the above experiments establish infants' sensitivity to prosodic characteristics at the word level. However, these results, obtained in simple discrimination tasks involving the presentation of only two contrasting stimuli, leave open the possibility that infants have only been reacting to very low-level acoustic differences.

Is it then the case that infants are able to extract from varied sets of speech stimuli those prosodic characteristics to which they are sensitive? One way to address this question is to study infants' ability to discriminate between two lists of phonetically varied stimuli, each list exemplifying a different facet of the property investigated. Given infants' ability to discriminate phonetic contrasts (see Jusczyk, 1997, for a review), discrimination in this paradigm is assumed to indicate two things: First, infants have extracted, across the phonetic variability, the property shared by the stimuli within a list; second, they used this stored information to contrast the two lists (see Bijeljac-Babic, Bertoncini & Mehler, 1993, for a discussion of this assumption). Using a similar procedure, Sansavini, Bertoncini & Giovanelli (1997) established that Italian newborns extract the accentual pattern of Italian words.

In the present study, we investigate whether newborns are sensitive to the pitch contour of words and whether they can extract this information from lists of words. Specifically, we tested French-born newborns' ability to discriminate two lists of Japanese phonetically varied bisyllabic words that differ in pitch contour (descending versus ascending). Japanese was chosen because pitch contour is lexically-specified in this language, resulting from the succession of the high- and low-pitched morae[1] which constitute the words. Japanese also allowed for a perfect control of the phonetic inventory of the two lists in providing pairs of words made up of the same phonemes but different pitch contours (and meanings). Discrimination of the lists would suggest that newborns can extract the pitch contour characteristics of words, and use this information to compare lists of words.

A total of 121 healthy, full-term newborns were tested at the Baudelocque Maternity Hospital (Paris, France). Forty infants (15 girls and 25 boys) completed the experiment. When test took place, their mean age was 68 h (*SD* = 24). Eighty-one infants were excluded for the following reasons: falling asleep (34), rejecting the pacifier or crying (15), irregular or insufficient sucking (17), failing to achieve the familiarization criterion within 18 min (15).[2]

Two lists of 24 bisyllabic (bimoraic) Japanese words were recorded by a female native speaker and stored in a digital form on the computer running the experiment. One list consisted of High-Low (H-L) words and the other of phonetically identical Low-High (L-H) words. The words used were: /ame/, /eki/, /huri/, /hashi/, /iki/, /ishi/, /jiku/, /joshi/, /kaki/, /kame/, /kami/, /kiji/, /mane/, /nari/, /nomi/, /oki/, /roku/, /tashi/, /toki/, /toshi/, /umi/, /uri/, /waki/, /yuki/.

Mean duration, fundamental frequency and energy (root mean square) of the first and second syllables were calculated for the two lists of words (see Table 1). These analyses show significant differences in pitch contours between the two types of words. They also show that the decrease in energy between the two syllables, found for both types of words, is more pronounced for the H-L words than for the L-H words.

Infants were placed in a reclining seat located in a sound-attenuated room. They

TABLE 1
Acoustic analysis of the stimuli

| | | High-Low words | | | Low-High words | | | |
|---|---|---|---|---|---|---|---|---|
| | | *syll.1* | *syll.2* | $HL_{diff}$ | *syll.1* | *syll.2* | $LH_{diff}$ | $LH_{diff}$ - $HL_{diff}$ |
| F0 (Hz) | *M* | 199.8 | 141.0 | –58.7 | 155.2 | 194.4 | 39.2 | 97.9*** |
| | *SD* | 11.8 | 13.4 | 13.8 | 6.1 | 8.0 | 10.5 | 19.3 |
| duration (ms) | *M* | 176.9 | 258.4 | 81.5 | 179.6 | 265.9 | 86.3 | 4.9 |
| | *SD* | 53.4 | 42.2 | 78.8 | 52.6 | 40.1 | 80.9 | 43.8 |
| energy (RMS) | *M* | 60.2 | 50.7 | –9.5 | 57.9 | 54.0 | -3.9 | 5.6*** |
| | *SD* | 3.9 | 8.1 | 6.8 | 3.3 | 7.2 | 6.6 | 2.3 |

*** $p < .001$

were given a pacifier, held by a mechanical arm, which was connected to a computer via a pressure transducer and an analog-to-digital board. The computer detected each suck, saving its amplitude and time of occurrence. Sucks above a specified amplitude, whose value was the same for all subjects and corresponded to about 80% of all sucks, were considered of high-amplitude (HA). Even though the computer entirely controlled the experiment, additional precautions were taken to avoid bias from the experimenter. The experimenter sat out of sight and reach of the infants, and could only intervene to readjust the pacifier before the last 5 min of familiarization (their data being otherwise discarded).

The experiment started with a 2-min baseline during which the infants' sucking rates in silence were registered. The familiarization period then began, during which stimuli were randomly picked out of one of the lists and delivered contingently upon HA sucks via a stereo amplifier and two loudspeakers. However, if the infant sucked continuously, the maximal rate of stimulus presentation was of one stimulus every 1200 ms. The familiarization period lasted at least 5 min, ending when a 25% decrement in sucking rate over 2 consecutive minutes compared to the rate of the immediately preceding minute was reached. The list of stimuli was then changed for the infants in the experimental group, while infants in the control group continued to be presented with the same list. The test period lasted 4 min.

Ten subjects were randomly assigned to each of the 4 sub-groups resulting from the crossing of the factors Condition (control vs. experimental) and Order (L-H words first vs. H-L words first): HL-HL, LH-LH, HL-LH, LH-HL.

The average sucking rates of the experimental and control groups (collapsed across Order) are displayed in Figure 1. Several analyses of variance (ANOVAs) with the main factors of Condition and Order were conducted to verify that these factors did not affect infants' behavior before the switch. Neither factor affected the mean duration of the familiarization phase [$F(1, 36) < 1$ for both factors] and the HA sucking rates during baseline and the last 5 and 2 preshift minutes [Condition: $F(1, 36) = 1.23, p = .27$ for the baseline; $F(1, 36) < 1$ for the other 2 periods; Order: $F(1, 36) < 1$ for all periods].

To evaluate the effect of the change of stimuli, the last 2 preshift minutes were compared with the first 2 postshift minutes. These comparisons revealed a significant increase in sucking rates in the experimental group [$F(1, 19) = 9.87, p = .005$], and no difference in the control group [$F(1, 19) < 1$]. The significant interaction between Shift (2 last preshift min vs. 2 first postshift min) and Condition [$F$(1,

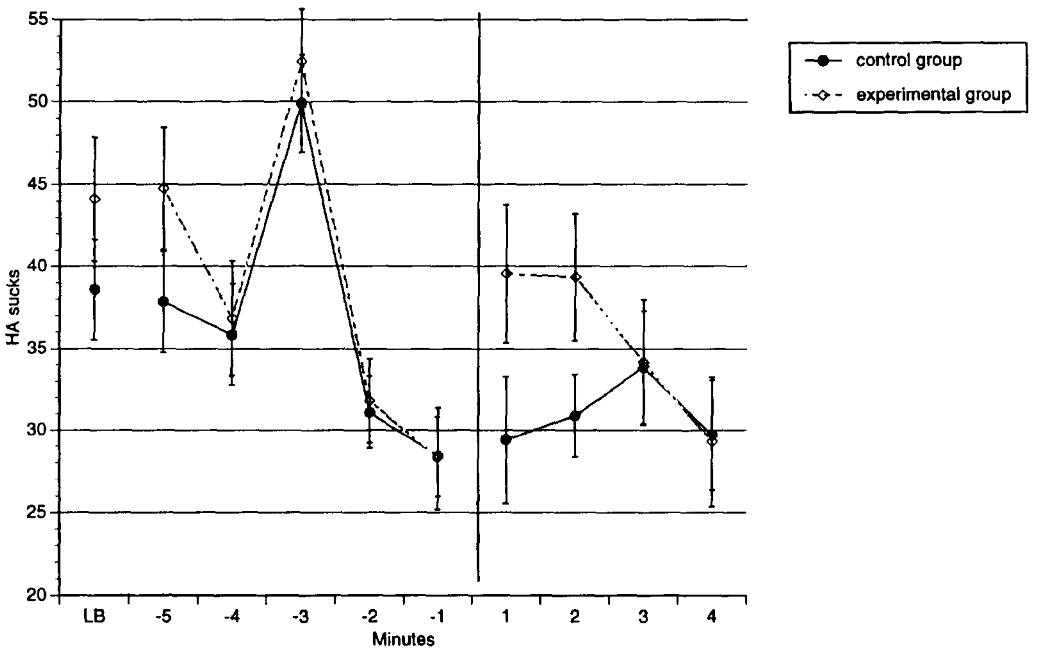


FIGURE 1
Sucking-rate averages during baseline (BL), last 5 min of familiarization (–5 to –1), and 4 min of test (+1 to +4) for the experimental and control groups.The bars above and below each point indicate the standard error of the mean.

38) = 5.15, $p$ = .03] indicated that sucking recovery was significantly larger in the experimental group than in the control group.[3] Moreover, the interaction between Condition, Shift and Order did not reach significance [$F(1, 36) < 1$], indicating that Order had not significantly influenced the way infants had reacted to the stimulus change.

This experiment shows that newborns can discriminate lists of bisyllabic Japanese words that differ in pitch contour, extending to newborns the earlier findings by Bull et al. (1985) and Karzon and Nicholas (1989). Moreover, given that infants were presented with phonetically varied word lists, discrimination is assumed here to reflect that newborns extracted the common pitch contour dimension of the words in each list. It is also assumed that their response was based on the comparison between stored information about the pitch contour of the familiarized list of words and that of the incoming words of the test phase.

Second, the fact that the subjects were born from French-speaking families suggests that this pitch sensitivity does not result from prenatal or brief postnatal exposure to speech, as the pitch dimension does not differentiate lexical entries in French (although similar but unsystematic variations can occur). We suggest that this sensitivity is part of the universal repertory of infants' innate abilities.

Finally, future research will have to address issues raised by our finding regarding early lexical acquisition and lexical processing by adults. First, if the existence of a correlation between ascending pitch contours and word boundaries in Japanese (suggested by an inspection of the data of Kuribayashi, 1996) is confirmed, do Japanese-learning infants initially use this regularity to extract words from fluent speech, similarly to the way Englishlearning infants use the lexical stress pattern regularities of English in early word segmentation (see Jusczyk, 1997, for a review)? Second, does the evolution of pitch sensitivity differ according to whether the language acquired uses pitch to contrast lexical entries, or does not? Such a phenomenon has been found for the sensitivity to another prosodic dimension: accentuation (Dupoux, Pallier, Sebastian & Mehler, 1997).

In conclusion, this study shows that newborns can discriminate lists of phonetically varied Japanese words that differ by their pitch contour. These results, together with those of Sansavini et al. (1997), suggest that newborns can extract and store prosodic information about the accentuation pattern and pitch contour of words.

**Acknowledgments:** This research was supported by grants from the Ministère de l'Enseignement Supérieur et de la Recherche to the first author and the Human Frontiers Science Program. We would like to thank Anne Cutler, Peter Jusczyk, Takashi Otake, Carolyn Rovee-Collier, Ruth Tincoff and two anonymous reviewers for helpful comments on a previous version of this paper.

## NOTES

Thierry Nazzi is now at University of California at Berkeley; Caroline Floccia is now at Université de Franche-Comté, Besançon; Josiane Bertoncini is now at Laboratoire Cognition et Développement, CNRS-Paris V.

1. The mora is the rhythmic unit of spoken Japanese and the written unit of Japanese *kana* orthographic system. Many morae correspond to syllables (V, CV), but there are 3 types of subsyllabic morae: vowel lengthening (as in CV-V), post-vocalic nasals (as in CV-N) and geminate consonant (as in CV-Q).
2. High attrition rates are common with the HA sucking procedure, due to newborns' unstable states of awakening. For example, to include 40 infants in the final analyses, 108 (van Ooijen et al., 1997), 91 and 92 (Nazzi et al., 1998, Exp. 1 and 2) newborns were tested.
3. ANOVAs including the third and fourth minutes of test showed non-significant interactions between Shift and Condition, a usual result in this procedure which indicates that the infants' reaction to the change of stimuli is transient.

## REFERENCES

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

Bull, D., Eilers, R. E., & Oller, D. K. (1984). Infants' discrimination of intensity variation in multisyllabic stimuli. *Journal of the Acoustical Society of America, 76,* 13–17.

Bull, D., Eilers, R. E., & Oller, D. K. (1985). Infants' discrimination of final syllable fundamental frequency in multisyllabic stimuli. *Journal of the Acoustical Society of America, 77,* 289–295.

Dupoux, E., Pallier, C., Sebastian, N., & Mehler, J. (1997). A destressing "deafness" in French? *Journal of Memory and Language, 36,* 406–421.

Eilers, R. E., Bull, D. H., Oller, D. K., & Lewis, D. C. (1984). The discrimination of vowel duration by infants. *Journal of the Acoustical Society of America, 75,* 1213–1218.

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10,* 279–283.

Gleitman, L. R., & Wanner, E. (1982). Language acquisition: The state of the state of the art. In E. Wanner, & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 3–48). New York, NY: Cambridge University Press.

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

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

Karzon, R. G. (1985). Discrimination of polysyllabic sequences by one- to four-month-old infants. *Journal of Experimental Child Psychology, 39,* 326–342.

Karzon, R. G., & Nicholas, J. G. (1989). Syllabic pitch perception in 2- to 3-month-old infants. *Perception and Psychophysics, 45,* 10–14.

Kuribayashi, H. (1996). Gendai nihongo no akusento no kata no bunpu ? denshi bukku ban daijiri o siryoo to site (Distribution of the accent patterns of modern Japanese-on the basis of the electronic version of Daijirin). *Kenkyuu Kiyoo, 51,* 1–28.

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

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

Ooijen, B. van, Bertoncini, J., Sansavini, A., & Mehler, J. (1997). Do weak syllables count for newborns? *Journal of the Acoustical Society of America, 102,* 3735–3741.

Sansavini, A., Bertoncini, J., & Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental Psychology, 33,* 3–11.

**21 July 1997; Revised 08 April 1998** ■