INFANT BEHAVIOR AND DEVELOPMENT 20 (2), 1997, pp. 263–269 ISSN 0163-6383

BRIEF REPORT

Non-nutritive Sucking and Sentence Processing

SARAH HESKETH
Cognitive Development Unit–Medical Research Council, London

ANNE CHRISTOPHE
Cognitive Development Unit–Medical Research Council, London

GHISLAINE DEHAENE-LAMBERTZ
Laboratoire de Sciences Cognitives et Psycholinguistique, Paris

A variant of the non-nutritive sucking method was explored to test 2-month-old infants' perception of whole sentences. Three consecutive high-amplitude sucks were required to trigger one sentence, and each subject was submitted to two changes in stimulation, one experimental (language change) and one control (speaker change). The results were significant, showing that the procedure is appropriate for the study of language perception in 2-month-old infants.

non-nutritive sucking 2-month-old infants speech perception.

Research has shown that human infants are extremely well-equipped to process speech and acquire language. For instance, very young infants are able to perceive all the phonemes from the languages of the world (Eimas, Siqueland, Jusczyk, & Vigorito, 1971). They can also distinguish their mother tongue from a foreign language. This is a particularly crucial ability, since infants could not possibly learn a language (that is, discover the regularities shared by a number of sentences) if they worked on a database containing sentences from several different languages (Mehler, Dupoux, Nazzi, & Dehaene-Lambertz, 1996).

To summarize infants' abilities to distinguish between languages, it has been shown that newborns can discriminate between their mother tongue and a foreign language. Mehler et al. (1988) found that 4-day-old French infants discriminate between French (their mother tongue) and Russian stimuli. In addition, these infants are able to discriminate between utterances in two foreign and unfamiliar languages, namely English and Italian (see Mehler & Christophe, 1995, for a reanalysis of the original data; see also Nazzi, Bertoncini, & Mehler, in press, for a replication) In this same study, American infants of 2 months of age were shown to exhibit a different behavior. Although they were able to discriminate between English (their mother tongue) and Italian, they failed to show any recovery of interest when switched from one foreign language to another, in this instance, French and Russian. A possible interpretation of this counter-intuitive result is that 2-month-old infants have sufficient knowledge of their mother tongue to be able to filter out foreign languages as being not relevant, while newborns still attempt to analyse any speech sample they are exposed to. Congruent with this interpretation, 2-month-old infants show a special reaction to their mother tongue. American 2-month-olds orient faster to sentences in English than to sentences in French, while French 2-month-olds exhibit the reverse pattern (Dehaene-Lambertz & Houston, personal communication). Even newborns, though, display a special reaction to their mother tongue. Moon, Cooper and Fifer (1993) tested 2-day-old monolingual American infants, half of whom had Spanish-speaking parents and half English speaking. They observed that infants sucked more while listening to their mother's native language than while listening to the foreign language.

Most of these studies have been replicated successfully using low-pass filtered speech. This means that only frequencies below 400 Hz are used (which include the overall properties of speech, such as intonation and rhythm), thereby

Direct all correspondence to: Anne Christophe, Laboratoire de Sciences Cognitives et Psycholinguistique, MSH, 54 Bd Raspail, 75270 PARIS Cedex 6 FRANCE.

filtering out the higher-frequency components which carry phonemic information. It is therefore probable that infants' ability to discriminate between languages is based on a representation of speech prosody. However, we still know little about the nature of the prosodic representation that infants use to classify languages.

A variety of methods have been used in these studies to test infants' processing of whole sentences. Mehler et al. (1988) have used a standard habituation-recovery procedure of two behavioral responses depending of the age of the infants. They have measured the number of sucks produced while newborns were listening to speech passages (the sound presentation being independent of infants' sucking behavior) and the looking time to a picture placed in front of the loudspeaker in 2-month-old infants. Moon et al. (1993) also used the sucking response, but made the language presented (Spanish vs English) contingent on the duration of pauses between bursts (a relationship that infants failed to exploit in this particular experiment). Dehaene-Lambertz and Houston (personal communication) measured the latency to orient towards auditory stimulation in a setting where sentences from both languages were randomly presented to the right or left of the infant. Although it is reassuring to observe the same results replicated with such a variety of experimental methods, finer interpretations may need comparisons within the same paradigm across different ages.

The present study is an attempt at defining an appropriate paradigm. We set out to replicate a known result, namely the ability of 2-month-old infants to discriminate between their mother tongue (English) and a highly different language (Japanese). Due to the immaturity of the motor control system in infants and to the fact that the high-amplitude sucking procedure is one of the best studied infant methods (Floccia, Christophe, & Bertoncini, 1997; Trehub & Chang, 1977; Williams & Golenski, 1978), we chose to focus on this technique. It enables the assessment of discrimination between stimuli that are not preferred, and can be used with both newborns and 2-month-olds, where previous results suggest that developmental changes take place. However, the method needs to be adapted to the study of infants' perception of whole sentences (although Mandel, Jusczyk, & Kemler-Nelson, 1994, used it successfully with short sentences, 1.2 s on average). Indeed, the traditional non-nutritive sucking method relies on the delivery of one speech stimulus per suck. Since sucks occur typically every 500–600 ms within bursts, such a reinforcement pattern ensures a good correspondence between sucks and sounds when sound stimuli are short. However, the correspondence becomes tenuous with a stimulus duration of several seconds, as is typically the case for whole sentences. In the present study, therefore, we chose to preserve the contingent relationship between sucking and sound that has been shown to be crucial to the method (Floccia et al., 1997; Williams & Golenski, 1978) and to present one sentence whenever the infant made three high amplitude (HA) sucks at the beginning of a burst.

Furthermore, we measured both the sucking response produced by the infants and the sentences thereby triggered. Indeed, because of the contingent relationship between sucking and sound, the amount of sucking is assumed to provide an index of the infant's interest in the sound. The number of sound stimuli triggered is simply a more direct measure of the same thing. These measures are, of course, correlated (because of the contingent relationship between sucking and sound), but the correlation may not be perfect in the present situation where three consecutive HA sucks are required to trigger a sentence. In addition, while the high amplitude sucking (HAS) technique traditionally involves a between-subject comparison, we submitted each infant to two changes in stimulation, one experimental and one control. Although this procedure lengthens the experimental session, at the risk of losing more infants, it should provide more robust within-subject statistics.

Most studies contrasting different languages have used bilingual speakers to make sure that infants did not react to a change in speaker rather than to a change in language. However, recent studies have suggested that bilingual speakers do not behave like two independent monolinguals in fine-grained perception tasks, even though they were judged to be perfect in both their languages and had learned them both at a young age (Cutler, Mehler, Norris, & Segui, 1983, 1992). Instead, these bilinguals behaved like monolinguals for only one of their mother tongues (the so-called "dominant" language), while they behaved differently for the other language. Using only monolingual speakers, therefore, ensures that the spoken input will be fully congruent with the language. In such a design, a change in language necessarily involves a change in speaker, therefore, the appropriate control for this condition is a change of speaker within the same language. The rationale behind this comparison is that a change in language is more relevant for the infant's speech processing system than a change in speaker alone—as mentioned above, different languages must be kept apart in the process of language acquisition, or infants exposed to more than one language would get confused (Mehler et al., 1996). We thus expect the language change to be more interesting to infants than the speaker change. In order to maximize our chances to observe a differential response to language and speaker changes, we made the speaker change as inconspicuous as possible. Several speakers were used in each phase of the experiment, all females whose voices were perceived as similar by adult listeners.

The stimuli consisted of 80 sentences, half in English, half in Japanese, which had a syllabic length between 15 and 21. These were recorded by four female native English speakers and four female native Japanese speakers. Speakers were naive as to the aim of the experiment and were instructed to read as naturally as possible. Each sentence was digitized at 16 kHz and stored as an independent file on the experimental computer. Ten sentences from each speaker were selected and matched for syllabic length (M = 17.8 syllables) and duration (means of speakers ranged between 3.0 and 3.1 s).

Each infant underwent two changes in stimulation, one experimental (language) change, the other, a control (or speaker) change. Half the infants received the experimental change first and the control change second. In addition, the order of presentation of languages and of speakers was counterbalanced across subjects. This yielded eight conditions (see Table 1).

Subjects were seated in a car seat placed in a sound-proofed chamber. A standard (steam sterilized) pacifier was connected to the seat by way of a mechanical arm. One experimenter was seated out of view behind the infant and checked that the pacifier stayed in the infant's mouth throughout the experiment. The experimenter was blind to the experimental condition and listened to a masking tape over headphones for the duration of the experiment.[1] A plastic tube connected the pacifier to the computer outside the chamber via a pressure transducer. A second experimenter monitored the experiment on the computer. The computer recorded the pressure of the infant's sucks via an analogue-digital card (NIDAQ), detected the sucking responses on the basis of speed of increase and decrease in sucking and an amplitude threshold, and delivered the sentences through a ProAudio 16 sound board according to the reinforcement schedule (see below). The computer also saved both the moment and amplitude of each suck as well as the stimuli triggered by the sucks.

TABLE 1
Experimental Conditions[a]

	Phase 1	Phase 2	Phase 3
Language	Eng. 1 + 2	Jap. 1 + 2	Jap. 3 + 4
Change	Eng. 3 + 4	Jap. 3 + 4	Jap. 1 + 2
First	Jap. 1 + 2	Eng. 1 + 2	Eng. 3 + 4
	Jap. 3 + 4	Eng. 3 + 4	Eng. 1 + 2
Language	Eng. 1 + 2	Eng. 3 + 4	Jap. 3 + 4
Change	Eng. 3 + 4	Eng. 1 + 2	Jap. 1 + 2
Second	Jap. 1 + 2	Jap. 3 + 4	Eng. 3 + 4
	Jap. 3 + 4	Jap. 1 + 2	Eng. 1 + 2

Note: [a] Eng. 1+2: English sentences, speakers 1 and 2; Eng 3+4, English sentences, speakers 3 and 4; and so on.

The experiment started with a short baseline without stimulation (about 30 s). The first phase of the experiment then began, during which infants heard sentences in either English or Japanese (see Table 1) contingent upon their high-amplitude (HA) sucks. Pilot subjects run in this procedure showed that many 2-month-olds failed to start sucking on the pacifier and were consequently denied the opportunity to trigger a sentence (since three consecutive sucks were required). A "shaping" phase was therefore introduced in the present experiment to encourage infants to become interested in the experimental situation. For the first six sentences, only one high-amplitude suck was sufficient to trigger a sentence, increasing to two consecutive HA sucks to trigger the next six sentences. After this, three HA sucks were required to trigger each sentence (such that there was less than 1 s between two consecutive sucks). There was an

ISI of at least 400 ms between two consecutive sentences. Within each phase of the experiment, the order of presentation of the sentences was quasi-random for each infant (i.e., all sentences were presented once in random order, then reshuffled and presented once, and so on, to the end of the phase).

A switch in stimulation occurred after a predefined habituation criterion had been met. For 2 consecutive minutes, the infant's HA sucking rate had to be less than 80% of the maximum sucking rate from the beginning of the experiment (excluding the very first minute of stimulation). In addition, the first switch of stimulation could not occur until the infant had shown a sufficient level of activity (at least 20 HA sucks in at least 1 min). Finally, at least one sentence had to be triggered during the last minute before a switch in stimulation to avoid having a silent period between two phases. Some infants sucked at a very low amplitude and did not manage to trigger many sentences (nor reach the 20 sucks criterion). In those cases, the experiment was restarted, and the sensitivity for sucking detection was doubled. Each phase of the experiment lasted at least 5 min.

A total of 36 infants aged between 6–12 weeks were recruited from the Cognitive Development Unit database to participate in the study. Of these subjects, a total of 16 infants, mean age: 8 weeks, 6 days, successfully completed the experiment. The remaining 20 infants were excluded from the final analysis for the following reasons: failed to start sucking (9); completed one switch only (2); cried (5); sucked at low amplitude (4). Subjects were randomly assigned to one of the eight conditions prior to testing. The average length of a test session was 19.8 min (range: 15 to 31 min), and subjects heard on average 55 sentences during phase 1 (range: 17 to 108).

Both the number of sucks per minute and the number of sentences triggered per minute were analyzed. An analysis of infants' responses during phase 1, before any experimental manipulation, revealed that infants listening to Japanese sucked significantly more, $t(14) = 2.43$, $p < 0.03$), and heard more sentences, ($t(14) = 2.39$, $p < 0.04$) than those listening to English. This may either reflect a sampling bias (with only 8 infants in each subgroup, this being a between-subject comparison of absolute activity rates), or reveal more interest for the foreign language in this situation. To assess the effect of the experimental manipulation, two kinds of analyses were performed on the data: ANOVAs and nonparametric tests. The relevant parameter is the increase in response rate for a change in language, compared to the increase in response rate for a change in speaker only (within-subject). To measure these increases, we used 2 min before and after each switch in stimulation.

Table 2 presents the increases in response rate per subject, for the experimental (language) switch and for the control (speaker) switch. The difference between these two measures represents a discrimination index per infant: Whenever this value is positive, the infant reacted more to the language change than to the speaker change.

A Wilcoxon signed rank test showed that the median of the sucking discrimination index was significantly above 0 (6 values below 0, $M = 3.5$; 8 values above 0, $M = 10.5$; $Z = 2.0$, $p < .05$). The same result holds for the number of sentences triggered (1 value below 0, $M = 2.5$; 15 values above 0, $M = 8.9$; $Z = 3.4$, $p < 0.001$).

In the ANOVAs, the dependent measure was the dishabituation score for the Experimental and Control switches. There was one within-subject factor (Experimental vs. Control switch) and three between-subject counterbalancing factors, Order (experimental switch first, vs. control switch first), Language (English first vs. Japanese first) and Speaker (starting with speakers 1 and 2 vs. starting with speakers 3 and 4). For sucking rates, the analysis showed a significant effect of the Experimental factor, $F(1, 8) = 5.7$, $p < .05$); none of the counterbalancing factors had any significant main effect, nor did they interact with the Experimental factor. The same results held for the number of sentences triggered: main effect of the Experimental factor, $F(1, 8) = 12.7$, $p < .01$); no significant effect of any of the counterbalancing factors, no interactions between the Experimental and counterbalancing factors. In particular, there was no interaction between Language and the Experimental factor (both $F(1, 8) < 1$, for sucks and sentences), therefore, no asymmetry in the discrimination behavior induced by which language was presented first (infants being switched from English to Japanese did not react

TABLE 2
Increases in Response Rate per Subject for Experimental and Control Switches

Condition:	Start with:	Sucking Responses: Exp.	Cont.	Diff.	Sentences Triggered: Exp.	Cont.	Diff.
Language	Eng. 1 + 2	1	2	-1	1	0.5	0.5
		0.5	-7.5	8	0.5	-1	1.5
Change	Eng. 3 + 4	8	0	8	1.5	0	1.5
		24.5	-19	43	4	-3.5	7.5
First	Jap. 1 + 2	13.5	9	4.5	5	0	5
		11.5	14.5	-3	4	3	1
	Jap. 3 + 4	2.5	-18	20.5	1	-2.5	3.5
		-1.5	-0.5	-1	-0.5	0	-0.5
Language	Eng. 1 + 2	-0.5	0	-0.5	1	0	1
		12	-11	22.5	2	-2	4
Change	Eng. 3 + 4	10	12.5	-2.5	2.5	0.5	2
		25.5	25.5	0	5	4	1
Second	Jap. 1 + 2	-0.5	-22	21.5	0.5	-3	3.5
		-6	-6	0	-1	-1.5	0.5
	Jap. 3 + 4	2	-14	16	-1	-2	1
		9.5	11.5	-2	1	0.5	0.5
			M	8.4		*M*	2.09
			SD	12.6		*SD*	2.02

more nor less than infants being switched from Japanese to English).

In this experiment, we successfully replicated the known result that 2-month-old infants can distinguish between sentences from their mother tongue and sentences from a very different foreign language (in this case, English and Japanese). This indicates that the experimental technique we used is appropriate for studying the perception of long segments of speech in young infants. We tested this technique with a language discrimination design, but it can also be used to tackle other research questions, such as infants' perception of prosody within their own language. In addition, we extended this result to a multilocutor situation. Infants reacted significantly more to a change in language than to a change in speakers within the same language (even though they might have reacted more to a change in speaker than to no change at all). This feature is particularly desirable in that it avoids having to design a procedure to check that a bilingual speaker is equally proficient in both languages (when, in fact, it may be impossible to find perfectly balanced bilingual speakers, see Cutler et al., 1992). It also allows full freedom in the choice of languages to be compared, as the only requirement is a few monolingual speakers for any language included in the study.

The number of sentences triggered appears to be a "cleaner" measure than the number of sucking responses. As can be seen from Table 2, there is less variability between subjects, as well as within subjects. For this reason, the effect appears to be more consistent when it is measured with the number of sentences triggered rather than with the number of sucking responses. This is not a counter-intuitive observation, since the number of sentences triggered is the most direct measure of the interest of the infant towards the language (whatever means the infants used in order to trigger the sentences). We therefore recommend the use of this measure in future studies.

One may wonder how much was gained by using a within-subjects design. To assess whether statistical power was improved compared to a standard, between-subjects design, we reanalyzed the data as if we had submitted 32 subjects to one switch of stimulation each (rather than 16 subjects to 2 switches). The results of the ANOVAs were in fact extremely similar (*sucks*: $F(1, 16) = 6.60$, $p < .05$; *sentences*: $F(1, 16)$ 14.40, $p < 0.01$); but the nonparametric test (a Mann-Whitney on independent samples) yielded only a marginal effect for sucks ($Z = 1.75$, $p = 0.08$), while it was still significant for sentences ($Z = 2.93$, $p < 0.005$). It therefore seems that some statistical power was gained by the within-subjects design.[2] Finally, the within-subjects design may lend itself better to the study of infants who suffered pathological events. It avoids having to match infants on pathology, which is often very difficult to do. One may consider doing single-case studies where an infant with a pathology is compared to a group of controls.

Part of our interest for the sucking paradigm was that this technique can be used with both newborns and 2-month-old infants, a period when developmental changes seem to take place. Would our design also work for newborns? Nazzi et al. (in press) used the same procedure with French 3-day-old infants and the very same English and Japanese sentences. They found that only about one-third of the infants who completed one switch of stimulation managed to complete the second switch (most newborns did not maintain an adequate level of attention for more than 15–20 min). More importantly, using a standard between-subjects analysis, they observed that newborns reacted significantly more to a change in language than to a change in speaker within the same language. It is therefore valid to compare newborn and 2-month-old infants' language discrimination abilities with the present experimental design.

To conclude, the ability of very young infants to discriminate languages appears to be very robust and has been observed with a variety of experimental paradigms and languages. The variant of the non-nutritive sucking procedure described in this paper appears to be a very useful procedure to explore young infants' capacities to process continuous speech and not just to test discrimination of short speech segments such as syllables. We hope that it will be useful in the future and enable experimenters to acquire new insights into the way infants perceive and process languages.

FOOTNOTES

1. Initially, a music tape was used, but it was found that this was not an effective mask for the naturally produced sentences used in this experiment, although many kinds of music were tried. To be effective, such a mask would have had to be played too loud and would have been heard by the infant. We finally discovered that a "babble noise" was an extremely effective mask: We superimposed four continuous streams of experimental sentences, two in each language. This masking made it impossible for the experimenter to identify the language of the sentences being played to the infant.
2. In fact, 12 subjects may be enough in future use of this technique (this is the number typically used by Jusczyk's team in their non-nutritive sucking experiments, see e.g., Hohne & Jusczyk, 1994). All the analyses performed were also significant on the first 12 or last 12 infants tested.

AUTHORS' NOTES

The work reported in this article was made possible by a grant from the Human Frontiers Science Program as well as by the Human Capital and Mobility programme. We especially want to thank John Morton for his continued support, Peter Jusczyk for his advice in handling 2-month-old infants in a non-nutritive sucking procedure, and Sarah Minister for her help in recruiting and testing subjects.

REFERENCES

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1983). A language-specific comprehension strategy. *Nature, 304*, 159–160.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology, 24*, 381–410.

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

Floccia, C., Christophe, A., & Bertoncini, J. (1997). High-amplitude sucking and newborns: The quest for underlying mechanisms. *Journal of Experimental Child Psychology, 64*, 175–198.

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

Mandel, D.R., Jusczyk, P.W., & Kemler-Nelson, D.G. (1994). Does sentential prosody help infants organize and remember speech information? *Cognition, 53*, 155–180.

Mehler, J., & Christophe, A. (1995). Maturation and learning of language in the first year of life. In M.S. Gazzaniga (Ed.), *The cognitive neuro-*

sciences: A handbook for the field (pp. 943–954). Cambridge, MA: MIT Press.

Mehler, J., Dupoux, E., Nazzi, T., & Dehaene-Lambertz, G. (1996). Coping with linguistic diversity: The infant's viewpoint. In J.L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 101–116). Mahwah, NJ: Erlbaum .

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

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

Nazzi, T., Bertoncini, J., & Mehler, J. (in press). Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance.*

Trehub, S.E., & Chang, H.W. (1977). Speech as a reinforcing stimulation for infants. *Developmental Psychology, 13*, 170–171.

Williams, J.A., & Golenski, J. (1978). Infant speech sound discrimination: The effects of contingent versus noncontingent stimulus presentation. *Child Development, 49*, 213–217.

16 July 1996; Revised 4 November 1996 ■