Brain & Language 148 (2015) 25–36

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Electrophysiological evidence of statistical learning of long-distance dependencies in 8-month-old preterm and full-term infants

C. Kabdebon [a,b,c,*], M. Pena [d], M. Buiatti [a,b,c], G. Dehaene-Lambertz [a,b,c]

[a] INSERM, U992, Cognitive Neuroimaging Unit, F-91191 Gif/Yvette, France
[b] CEA, DSV/I2BM, NeuroSpin Center, F-91191 Gif/Yvette, France
[c] University Paris 11, F-91191 Orsay, France
[d] Escuela de Psicología, Pontificia Universidad Católica de Chile, Chile

A R T I C L E I N F O

Article history:
Accepted 9 March 2015
Available online 9 April 2015

Keywords:
Language
Infant
Statistics
Learning
Event-related potentials
Frequency tagging
Phase-locking
Prematurity

A B S T R A C T

Using electroencephalography, we examined 8-month-old infants' ability to discover a systematic dependency between the first and third syllables of successive words, concatenated into a monotonous speech stream, and to subsequently generalize this regularity to new items presented in isolation. Full-term and preterm infants, while exposed to the stream, displayed a significant entrainment (phase-locking) to the syllabic and word frequencies, demonstrating that they were sensitive to the word unit. The acquisition of the systematic dependency defining words was confirmed by the significantly different neural responses to *rule-words* and *part-words* subsequently presented during the test phase. Finally, we observed a correlation between syllabic entrainment during learning and the difference in phase coherence between the test conditions (*rule-words* vs *part-words*) suggesting that temporal processing of the syllable unit might be crucial in linguistic learning. No group difference was observed suggesting that non-adjacent statistical computations are already robust at 8 months, even in preterm infants, and thus develop during the first year of life, earlier than expected from behavioral studies.

1. Introduction

Statistical learning is one of the most successful approaches to provide an efficient account of how infants learn the particular characteristics of their native language. From the distributional pattern of sounds in their environment that shapes their discrimination abilities (Maye, Weiss, & Aslin, 2008; Maye, Werker, & Gerken, 2002) to the co-occurrence of specific acoustic properties (low intensity, short duration, weak stress, particular distribution of phonemes) and reproducible positions in prosodic domains that help them to infer word category (open vs closed-class words) (Hochmann, Endress, & Mehler, 2010; Shi, Cutler, Werker, & Cruickshank, 2006), infants display particularly efficient capacities for analyzing the statistical properties of events in the speech signal from an early age on.

Infants are also rapidly able to keep track of the statistical dependency between two events. As transitional probabilities (TP) between adjacent syllables tend to be higher within words than between words, it has been hypothesized that infants could segment words from fluent speech using this difference between word-internal and word-external TP. Several studies confirmed that infants can use this strategy (Goodsitt, Morgan, & Kuhl, 1993; Johnson & Tyler, 2010; Saffran, Aslin, & Newport, 1996). Tracking adjacent probabilities backwards was similarly proposed to infer a word category given the preceding word (e.g. noun vs verb depending on the preceding determiner vs personal pronoun), a capacity reported in 12-month-olds (Mintz, 2003) and 14-month-olds (Höhle, Weissenborn, Kiefer, Schulz, & Schmitz, 2004; van Heugten & Shi, 2010).

However, linguistic structures are not limited to adjacent syllables. The syntactic structure of a language in particular relies on building dependencies between distant parts of a sentence, as for example in subject-verb agreement. However, tracking long-distance statistics requires on the one hand more complex computations to bypass adjacent dependencies and on the other hand larger memory buffers to discover them. Yet, once long-distance dependencies are acquired, the coding of long utterances is simplified and the memory load reduced. The advantage of using this strategy to discover the higher levels of the syntactic hierarchy is obvious, and we may wonder at what age infants can rely on this capacity to process speech. Thereafter, we will refer to this

* Corresponding author at: Laboratoire de Neuroimagerie Cognitive, INSERM, U992, CEA/SAC/DSV/DRM/NeuroSpin, Bat 145, point courrier 156, F-91191 Gif/Yvette, France. Fax: +33 1 69 08 79 73.
E-mail address: claire.kabdebon@gmail.com (C. Kabdebon).

http://dx.doi.org/10.1016/j.bandl.2015.03.005
0093-934X/

long-distance statistical dependency as AxC, i.e. syllable A predicts syllable C, independently of the intervening and variable x. AxC computations were first reported in 18-month-old infants who accepted sentences of the type "*is* runn*ing*" but not of the type "*can* runn*ing*", whereas 15 month-olds were not able to differentiate the two types of sentences (Santelmann & Jusczyk, 1998). Using simpler artificial languages but brief exposure of a few minutes, Gómez and Maye (2005) reported success in non-adjacent statistical learning at the younger age of 15 months but failure at 12, and in a very recent experiment, Marchetto and Bonatti (2014) reported an even younger age for the onset of this capacity. They showed that 12- and 7-month-olds were able to discover the systematic dependency between the first and third syllable of trisyllabic words (AxC words) embedded in an artificial stream of syllables. However, only the 12 month-olds were able to generalize the structure (A predicts C) to new words comprising a novel syllable between A and C.

A few EEG studies have brought additional evidence for an early ability to track long-distance dependencies. For example, in French, long-distance computations are crucial to notice that the sentence "*Je la donne*" (I give it) is grammatical, while "*Je la fraise*" (I strawberry it) is not. In both cases, the local transitions are correct because the function word "*la*" can be either a determiner or a clitic (thus "*je la*", "*la donne*" and "*la fraise*" are correct pairs of words). Yet, 24-month- French infants reacted to the ungrammatical sentences embedded in short stories by a positive electrical component over the parietal areas revealing that they were able to take into account long-distance dependencies during on-line processing of natural speech (Bernal, Dehaene-Lambertz, Millotte, & Christophe, 2010). At a much younger age, 4-months, Friederici, Mueller, and Oberecker (2011) also observed a similar positive component over centro-parietal regions in response to non-grammatical sequences. German infants who were exposed to 18 min of naturally produced Italian sentences (*Il fratello/La sorella sta x-ando* or *puo x-are*), progressively noticed the verb-inflexion exchange (*sta x-are* or *puo x-ando*) presented during short test periods regularly spaced during exposure. Prosodic cues might have here played a crucial role to help chunk the stream in smaller units in order to memorize the two edges of the second chunk (*sta x-are* or *puo x-ando*). This result at this age is remarkable given that the intervening x element was chosen among 32 possibilities, certainly overwhelming infant's capacity to memorize each AxC triplet. The verb-inflexion exchange was thus only noticeable if infants had kept track of the exact relation between A and C. In a final study (Mueller, Friederici, & Mannel, 2012), 3-month-old infants were exposed to isolated trisyllabic non-words. The words belonged to two AxC families, the intervening x syllable being drawn from a set of 20 syllables. Two types of deviants were randomly interspersed: the last syllable of the deviant words was either exchanged between the two families (rule violation) or presented a change of pitch (control violation). The group of infants as a whole did not show a significant mismatch response (MMR) to the exchanged syllable (rule violation). However, once the infants were sorted as a function of gender and of polarity of their MMR to pitch, a complex interaction between gender and polarity of the mismatch response was reported, indicating that male and female infants with a positive MMR to pitch showed no MMR to rule violation, while within the negative MMR pitch group, males showed a positive MMR to rule violation and females showed a negative one. This interaction was interpreted by the authors as resulting from a maturational advantage of girls over boys in rule learning.

To sum up this rapid review, contrary to the computation of transition probabilities between adjacent syllables which is robust from, at least, the second semester of life, the age of success for long-distance computations appears to be highly variable across studies. There is converging evidence between behavioral and EEG studies that toddlers are able to compute non-adjacent probabilities after 18 months, but the results in younger infants are less stable. However, whereas infants' capacities to compute adjacent transitional probabilities have been tested with close paradigms across ages, the complexity of the learning utterances widely differs across the studies presented above because of the size (one or two syllables) and variability of the intervening non-pertinent element (i.e. the pool of x), the number of A_C families to learn and thus the relative weight of the adjacent (Ax and xC) and non-adjacent (A_C) dependencies, the length of the sentence to analyze, the length of the training, etc. Thus our goal was here to confirm whether infants in the first year of life were able to compute long-distance dependencies, or whether there is a sizeable age gap between adjacent and non-adjacent statistical computations which would suggest that they may depend on different neural bases.

We tested healthy 8-month-old infants because, at this age, infants listening to long utterances display unquestionable sensitivity to its segmental properties and to the transitional probabilities between adjacent syllables (Saffran et al., 1996). Our participants belonged to a group of full-terms or to one of two possible groups of preterm infants, one evaluated at 8 months after birth and the other at 8 months after the age they should have been born (term age), to cover a wide range of maturation and exposure to the ex-utero environment and examine the influence of these factors on the emergence of this capacity. This study is part of a program of experiments designed to investigate the relative influence of neural maturation and exposure to a linguistic environment on the development of linguistic abilities. Some abilities may develop following a maturational calendar whatever the environment, while other abilities may depend on the duration of exposure to broadcast speech and social interaction. In that case, preterm infants should benefit from their longer ex-utero experience. Eventually, some abilities might be systematically delayed in preterm infants due to a non-optimal early environment, even when no brain lesion is observed. We already observed that neural maturation is crucial in the loss of electrophysiological responses to non-native linguistic features (Pena, Pittaluga, & Mehler, 2010; Pena, Werker, & Dehaene-Lambertz, 2012), whereas duration of ex-utero exposure is the key factor for the learning of the phonotactic rules of the native language (Gonzalez-Gomez & Nazzi, 2012), and for gaze-following (Pena, Arias, & Dehaene-Lambertz, 2014). Note, however, that if the capacity develops earlier than the tested age, results will be similar in all three groups.

To be as close as possible to the experimental paradigms used to test adjacent statistical computations (Saffran et al., 1996), we exposed infants to 2 min of an artificial monotonous stream of concatenated syllables in which trisyllabic words were embedded, separated by a 25 ms subliminal pause. The words belonged to three families, each one characterized by an A_C dependency, with the intermediate syllable x coming from a pool of 3 syllables (Table 1). In the stream, each pair of adjacent syllables (Ax, xC, CA', etc.) had a similar range of low transition probability, whereas transition probabilities between non-adjacent syllables structured the continuous stream into trisyllabic words. Trisyllabic items were subsequently presented in isolation during a test part, and were consistent, or not, with the structure of the stream.

This paradigm was successfully tested in adults (Pena, Bonatti, Nespor, & Mehler, 2002). Adults exposed to such AxC stream and subsequently presented with isolated trisyllabic items identified as words the triplets consistent with the A_C dependency, although they never appeared in the stream (thereafter *rule-words*), but not the xca' triplet (*part-words*) that did appear in the