



Prosodic grouping at birth

Nawal Abboub^{a,*}, Thierry Nazzi^{a,b}, Judit Gervain^{a,b}

^aLaboratoire Psychologie de la Perception, Université Paris Descartes, Sorbonne Paris Cité, Paris, France

^bLaboratoire Psychologie de la Perception, CNRS, Paris, France



ARTICLE INFO

Article history:

Received 16 June 2015

Revised 26 July 2016

Accepted 6 August 2016

Keywords:

Newborn infants

Perceptual biases

Prosodic grouping

Prenatal exposure

Bilingualism

Near-infrared spectroscopy

ABSTRACT

Experience with spoken language starts prenatally, as hearing becomes operational during the second half of gestation. While maternal tissues filter out many aspects of speech, they readily transmit speech prosody and rhythm. These properties of the speech signal then play a central role in early language acquisition. In this study, we ask how the newborn brain uses variation in duration, pitch and intensity (the three acoustic cues that carry prosodic information in speech) to group sounds. In four near-infrared spectroscopy studies (NIRS), we demonstrate that perceptual biases governing how sound sequences are perceived and organized are present in newborns from monolingual and bilingual language backgrounds. Importantly, however, these prosodic biases are present only for acoustic patterns found in the prosody of their native languages. These findings advance our understanding of how prenatal language experience lays the foundations for language development.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Learning about language depends critically on a complex interplay between neurobiologically constrained processing mechanisms, perceptual biases and linguistic input. At birth, infants possess many language-general abilities. They can discriminate between most speech sounds (Cheour-Luhtanen et al., 1995; Werker & Gervain, 2013); and between rhythmically different languages they never heard before (Nazzi, Bertoincini, & Mehler, 1998; Nazzi & Ramus, 2003). Moreover, they prefer speech over a variety of non-linguistic sounds (Decasper & Spence, 1986; Vouloumanos & Werker, 2007) and infant- over adult-directed speech (Fernald & Kuhl, 1987). However, as hearing is operational from the 24th to the 28th week of gestation (Hepper & Shahidullah, 1994), experience with spoken language starts in the womb, and some evidence of prenatal learning is found at birth. Indeed, newborns prefer their mother's voice over other female voices (Decasper & Fifer, 1980), their native language over a rhythmically different unfamiliar language (Mehler et al., 1988; Moon, Panneton Cooper, & Fifer, 1993), and their communicative cries reflect the prosody of the language they heard in utero (Mampe, Friederici, Christophe, & Wermke, 2009). Moreover, it has been shown that newborns who received bilingual prenatal exposure recognize both languages as familiar and can discriminate them from a

rhythmically different unfamiliar language (Byers-Heinlein, Burns, & Werker, 2010). Additionally, newborns are able to recognize stories heard during pregnancy (Decasper & Spence, 1986) or melodies to which they were exposed prenatally (DeCasper, 1994; Granier-Deferre, Bassereau, Ribeiro, Jacquet, & Decasper, 2011). Taken together, these findings constitute evidence that infants start learning about language while still in the womb, and that speech heard in utero has a more important impact on the development of speech perception and language learning than hitherto believed.

Speech experienced in utero, however, is different from broadcast speech transmitted through the air. Maternal tissues act as a low-pass filter, mainly transmitting sounds below 300–400 Hz (Gerhardt et al., 1992; Querleu, Renard, Versyp, Paris-Delrue, & Crèpin, 1988). As a consequence, prosody, the global melody and rhythm of speech, is relatively well preserved and transmitted to the fetal inner ear, whereas more detailed, phonetic aspects are disrupted (Querleu et al., 1988). Importantly, prosody is a powerful cue that infants have been shown to make use of during language acquisition. For instance, newborns rely on prosody to discriminate languages (Nazzi, Bertoincini, et al., 1998; Nazzi & Ramus, 2003), to detect boundaries in speech (Christophe, Dupoux, Bertoincini, & Mehler, 1994), differences in the pitch contour or lexical stress pattern of words (Nazzi, Floccia, & Bertoincini, 1998; Sansavini, Bertoincini, & Giovanelli, 1997) or even between function words and content words (Shi, Werker, & Morgan, 1999), on the basis of their different acoustic characteristics. They also use prosody to segment words out of the continuous speech stream (Johnson &

* Corresponding author at: Laboratoire Psychologie de la Perception (UMR 8242), CNRS & Université Paris Descartes, 45, rue des Saints Pères, Paris 75006, France.

E-mail address: n.abboub@gmail.com (N. Abboub).

Jusczyk, 2001; Jusczyk, Houston, & Newsome, 1999; Kooijman, Hagoort, & Cutler, 2009; Mattys, Jusczyk, Luce, & Morgan, 1999; Nazzi, Iakimova, Bertocini, Fredonie, & Alcantara, 2006; Nishibayashi, Goyet, & Nazzi, 2015) or to learn about the syntactic features of their native language (Hirsh-Pasek et al., 1987), such as its basic word order (Gervain & Werker, 2013; Nespor et al., 2008) or argument structure (Christophe, Gout, Peperkamp, & Morgan, 2003).

Thus, the variations in pitch, intensity or duration that carry prosody in the speech signal serve as robust and particularly important cues to language learning. Yet, how infants perceive these three acoustic dimensions at birth has remained largely unexplored, and whether language experience shapes the perception of these acoustic cues is currently heatedly debated. One issue at stake is the origin and developmental trajectory of the prosodic grouping bias known as the Iambic-Trochaic Law (ITL). Some authors have argued that the ITL is language-independent. Specifically, it has been claimed that the auditory system automatically groups sequences of sounds that differ in duration with the longest element in final position (i.e., prominence-final or iambic grouping), and sequences of sounds that differ in intensity or pitch with the loudest or highest element in initial position (i.e., prominence-initial or trochaic grouping). The ITL was initially proposed to explain the grouping of musical or non-linguistic sequences (Bolton, 1894; Cooper & Meyer, 1960; Woodrow, 1951). As a well-known example, people tend to perceive the fire truck siren as a sequence of two paired sounds, the first one being higher than the second one. This grouping principle was later extended to account for regularities in speech production and biases in speech perception in adults (Bion, Benavides-varela, & Nespor, 2011; Hay & Diehl, 2007; Hayes, 1995; Nespor et al., 2008). The proposal that the ITL is language-general is supported by studies showing that adult speakers of prosodically and rhythmically different languages such as English and French show similar grouping preferences (Hay & Diehl, 2007). Moreover, trochaic grouping on the basis of a pitch contrast was found in Italian adults, in Italian and French infants, whose native language makes little use of pitch cues in its prosody (Abboub, Boll-Avetisyan, Bhatara, Hoehle, & Nazzi, 2016; Bion et al., 2011), as well as in rats (de la Mora, Nespor, & Toro, 2013), suggesting not only that prosodic grouping preferences might exist in the absence of language experience, but also that they might be shared by humans and other mammals.

However, a recent alternative hypothesis has emerged, according to which prosodic grouping biases might, at least in part, be influenced by language experience. Supporting this view, recent cross-linguistic research has shown that although English and Japanese adults group sequences varying in intensity trochaically, only English, but not Japanese, adults group sequences varying in duration iambically (Iversen, Patel, & Ohgushi, 2008). The two languages differ at the phrasal level, since Japanese has a trochaic rhythm (^Tokyo ni, Tokyo to, 'to Tokyo', with prosodic prominence marked by higher pitch on the content word 'Tokyo' in initial position; Gervain & Werker, 2013), whereas English has an iambic rhythm (to Ro:me, with prosodic prominence marked by lengthened duration on the content word 'Rome' in final position). Relatedly, while both German and French adults follow the ITL when presented with complex linguistic stimuli varying in intensity or duration, they nevertheless exhibit language-specific differences, German adults showing stronger ITL effects; moreover, effects based on pitch were found for German but not French adults (Bhatara, Boll-Avetisyan, Unger, Nazzi, & Höhle, 2013). Similar findings were found using complex non-linguistic stimuli (Bhatara, Boll-Avetisyan, Agus, Höhle, & Nazzi, 2015). The authors argue that these cross-linguistic differences reflect the fact that German has a predominantly trochaic word-level stress pattern, while French does not. Additionally, French is iambic at the phrasal

level, whereas German can have both rhythmic patterns. In infants, Japanese- and English-learning 7–8-month-olds (Yoshida et al., 2010) revealed a pattern of results similar to the one found in adults (Iversen et al., 2008) and bilingual Spanish and Basque 9–10-month-olds (Molnar, Lallier, & Carreiras, 2014) also showed consistent grouping for intensity, but not for duration. These early cross-linguistic differences were found essentially for duration, suggesting first, that the language environment might influence grouping preferences early on, and second, that the three acoustics cues are not affected in the same way by this cross linguistic modulation. This raises the question of how and when during development language experience starts modulating perceptual grouping biases.

Contributing to this debate, the current study will explore whether newborns already possess general perceptual mechanisms to group sounds according to prosodic cues, and whether these abilities are already modulated by the native language(s) heard in utero. If such perceptual biases are present early in development, they have the potential to help infants break into language.

A related point concerns the cerebral basis of prosodic grouping, which remains, to a large extent, unexplored. In adults, language comprehension, including morphosyntactic and semantic processing, is predominantly lateralized to the left hemisphere, while prosodic processing typically recruits a more dynamic network in the right hemisphere (Friederici, 2012; Hickok & Poeppel, 2007), although the lateralization of prosodic processing also depends on the functional relevance of prosody in the language studied and on the context. Left dominance may be observed if the prosodic cue used is lexically or morphosyntactically relevant such as lexical tone in adults who speak a tonal language (Gandour et al., 2004; Kreitewolf, Friederici, & von Kriegstein, 2014; Sato, Sogabe, & Mazuka, 2007, 2010). In infants, few studies have investigated the neural basis of prosodic processing in general, and none have specifically looked at how prosodic grouping is processed across these three acoustic dimensions. The few existing optical imaging studies investigating prosodic processing in general reported that sleeping neonates and 3-month-olds showed a right hemispheric specialization (Hamae, Watanabe, Nakano, & Taga, 2007; Sato et al., 2010; Telkemeyer et al., 2009), as do 4-year-olds (Wartenburger et al., 2007). Nevertheless in these studies, sentential prosody was tested in its full acoustic complexity. Thus it is still unclear how and where in the brain prosodic cues in isolation (i.e. variations in duration only, intensity only or pitch only) are processed and grouped. More specifically, we do not know whether, and if yes, how grouping on the basis of a single acoustic cue is perceived and processed in the developing brain.

The current study therefore sought to answer two questions. First, we explored the earliest foundations of the crucial ability to detect and process prosodic patterns. In particular, no study has as yet tested newborns' prosodic grouping biases and their neural correlates, a gap that the present study intends to fill. Accordingly, we tested prosodic patterns that vary along one of the three acoustic dimensions characterizing speech prosody: duration, intensity, and pitch. To do so, we used near-infrared spectroscopy (NIRS), an optical imaging technique ideally suited to test the youngest developmental populations (high motion tolerance, easy application, no carrier substance or magnetic field, no noise, etc.). NIRS has the advantage of providing good spatial localization, allowing us to identify the brain areas responsible for prosodic grouping. This technique has been widely used to explore the neural correlates of speech perception and language acquisition in newborns and young infants (Gervain, Berent, & Werker, 2012; Gervain, Macagno, Cogoi, Peña, & Mehler, 2008; Gomez et al., 2014; May, Byers-Heinlein, Gervain, & Werker, 2011; Peña et al., 2003; Telkemeyer et al., 2009). French, the language that our monolin-

Download English Version:

<https://daneshyari.com/en/article/7283751>

Download Persian Version:

<https://daneshyari.com/article/7283751>

[Daneshyari.com](https://daneshyari.com)